

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 01-16-03
Searcher: Beverly C 4994
Terminal time: 25
Elapsed time: _____
CPU time: _____
Total time: 25
Number of Searches: _____
Number of Databases: 1

Search Site

_____ STIC
_____ CM-1
_____ Pre-S

Type of Search

_____ N.A. Sequence
_____ A.A. Sequence
_____ Structure
_____ Bibliographic

Vendors

_____ IG Suite
_____ STN
_____ Dialog
_____ APS
_____ Geninfo
_____ SDC
_____ DARC/Questel
✓ _____ Other CGN

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 08:51:46 ; Search time 57 Seconds
(without alignments)
1028.601 Million cell updates/sec

Title: US-09-830-111a-2

Perfect score: 2194

Sequence: 1 MASPALRIRSSRSIASLR.....ESPAKALEQLTKVLTISR 440

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: A_Geneseq_101002.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2194	100.0	440	22	AA1980.DAT.*
2	1243	56.7	451	23	AA1981.DAT.*
3	1064	48.5	502	23	AA1982.DAT.*
4	1028.5	46.9	537	23	AA1997
5	999	45.5	425	23	AA14998
6	996	45.4	403	23	AA14999
7	902.5	41.1	378	23	AA14999
8	863	39.3	377	18	AA130762
9	742	33.8	430	22	AA198466
10	658	30.0	321	21	AA19412

11	658	30.0	321	23	AB181705
12	658	30.0	321	23	AB181705
13	628.5	28.6	272	23	AB181763
14	561	25.6	297	21	AB19411
15	561	25.6	297	23	AB181704
16	561	25.6	297	23	AA172774
17	451.5	20.6	323	22	AA134728
18	451.5	20.6	323	22	AA138141
19	451.5	20.6	323	23	AA181873
20	445.5	20.3	329	22	AA135549
21	445.5	20.3	329	23	AA181874
22	445.5	20.3	333	20	AA124397
23	445.5	20.3	333	20	AA124398
24	439.5	20.0	335	21	AA121095
25	410	18.7	323	17	AA192060
26	409	18.6	320	19	AA174722
27	407.5	18.6	448	22	AB165155
28	405	18.5	322	22	AA136438
29	404.5	18.4	325	23	AA121871
30	402.5	18.3	325	20	AA101634
31	398.5	18.2	333	22	AA175107
32	397.5	18.1	333	23	AA121866
33	396.5	18.1	410	19	AA153921
34	395.5	18.0	337	23	AA121867
35	395	18.0	315	19	AA153922
36	395	18.0	315	23	AA121868
37	395	18.0	348	19	AA153920
38	393.5	17.9	320	23	AA121870
39	378.5	17.3	348	22	AA151415
40	371.5	16.9	322	22	AA137010
41	368	16.8	319	23	AB154681
42	368	16.8	538	21	AA181868
43	367.5	16.8	312	22	AA138119
44	362.5	16.5	321	23	AB149219
45	357	16.3	326	23	AB149754

ALIGNMENTS

RESULT 1

AA174623
ID AA174623 standard; Protein; 440 AA.

AC AA174623;

XX 23-MAY-2001 (first entry)

DT 23-MAY-2001 (first entry)

DE Saitoella complicata decaprenyl diphosphate synthase protein SEQ.2.

XX Saitoella complicata; decaprenyl diphosphate synthase; coenzyme Q10;

XX fungus.

OS Saitoella complicata.

XX WO200114567-A1.

XX 01-MAR-2001.

XX 24-AUG-2000; 2000WO-JP05659.

XX 24-AUG-1999; 99JP-0237561.

XX (KANF) KANEKA CORP.

XX Matsuda H, Kawamukai M, Yajima K, Ikenaka Y, Hasegawa J;

XX Takahashi S;

XX WPI; 2001-202937/20.

XX N-PSDB; AA174893.

XX DNA encoding a protein having decaprenyl diphosphate synthase activity

PT and microorganism for producing coenzyme Q10 -

XX PS Claim 2, Page 28-30; 32pp; Japanese.
XX CC The present invention describes a method for microbiologically producing
CC coenzyme Q10 at a high efficiency by using a gene of the synthesis of
CC coenzyme Q10 side chain originating in a fungus belonging to the genus
CC *Saitoella*. The present sequence represents the specifically claimed
CC *Saitoella* complicated protein having decaprenyl diphosphate synthase
CC activity. The protein having decaprenyl diphosphate synthase activity
CC can be used for producing coenzyme Q10.
XX CC
SQ Sequence 440 AA;
Query Match 100.0%; Score 2194; DB 22; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.9e-202;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASPARIRSSISSTASLSVTLRRTASAPSLRRCPTSPSSSMAAASASRLVEPD 60
DB 1 MASPARIRSSISSTASLSVTLRRTASAPSLRRCPTSPSSSMAAASASRLVEPD 60
QY 61 PNOPLINPLLVGPEMSNLTSNIRSLGSGHPSLDTVAKYVYVSGKHIRPLVLMQAQ 120
DB 61 PNOPLINPLLVGPEMSNLTSNIRSLGSGHPSLDTVAKYVYVSGKHIRPLVLMQAQ 120
QY 121 TEVAPKVOGKEKVEVPVNEGSLAPPEVLNDKNDPMNMNRSGPLTKGCEIRGQTSNIIASQ 180
DB 121 TEVAPKVOGKEKVEVPVNEGSLAPPEVLNDKNDPMNMNRSGPLTKGCEIRGQTSNIIASQ 180
QY 181 RLRLAETEMHHAASLHDDVIDASETRRNAPSGNOAFGNKALTAGPFLGRASVALARL 240
DB 181 RLRLAETEMHHAASLHDDVIDASETRRNAPSGNOAFGNKALTAGPFLGRASVALARL 240
QY 241 RNEVEIELLTVIANLVGEFEMQKNTVDAIEATATQETFPDYVLYQKTYLKTASLIKSC 300
DB 241 RNEVEIELLTVIANLVGEFEMQKNTVDAIEATATQETFPDYVLYQKTYLKTASLIKSC 300
QY 301 RASALLGATPEVADAAYAGRNIGLAFQIVDDMLDYTVSATDILGKPGADLQIGLATAP 360
DB 301 RASALLGATPEVADAAYAGRNIGLAFQIVDDMLDYTVSATDILGKPGADLQIGLATAP 360
QY 361 ALPAWKHAELGPMIKRKFSDGVERARELVESDGLKTRALAEYAKALDAITFP 420
DB 361 ALPAWKHAELGPMIKRKFSDGVERARELVESDGLKTRALAEYAKALDAITFP 420
QY 421 ESPARKALEQLTDKVLTRSR 440
DB 421 ESPARKALEQLTDKVLTRSR 440
RESULT 2
ABB83793
ID ABB83793 standard; Protein; 451 AA.
XX AC ABB83793;
XX DT 29-AUG-2002 (first entry)
XX DE Fungal decaprenyl diphosphate synthase SEQ ID NO 4.
XX KM Fungi; *Aspergillus*; *Leucosporidium*; coenzyme Q10; enzyme;
XX KW decaprenyl diphosphate synthase; enzyme.
XX OS *Aspergillus clavatus*.
XX FN WO200252017-A1.
XX PD 04-JUL-2002.
XX PF 27-DEC-2001; 2001MO-JP11523.
XX PR 27-DEC-2000; 2000JP-0398658.
XX

PA (KANF) KANEXA CORP.
XX PI Matsuda H, Kawamukai M, Yajima K, Ikenaka Y;
XX DR WPI; 2002-500767/53.
XX DR N-PSDB; ABB81605.
XX PT Microbial production of coenzyme Q10 by transformants transferred with
PT fungal decaprenyl diphosphate synthase gene, on industrial scale for
PT application in drugs
XX PS
PS Claim 4; Page 37-40; 52pp; Japanese.
XX CC The invention relates to a DNA sequence (I, ABB81604-ABB81605),
CC containing a sequence based on that of (I) but with some bases deleted,
CC added, inserted and/or substituted and encoding a protein
CC (ABB83792-ABB83793) with decaprenyl diphosphate synthase activity or
CC hybridisable with (I) under stringent conditions and encoding a protein
CC with decaprenyl diphosphate synthase activity. The DNA is useful for the
CC production of coenzyme Q10 for application in drugs. The coenzyme is
CC simply produced on industrial scale, efficiently, with significantly
CC productivity.
XX CC
SQ Sequence 451 AA;
Query Match 56.7%; Score 1243; DB 23; Length 451;
Best Local Similarity 57.2%; Pred. No. 1.2e-110;
Matches 261; Conservative 61; Mismatches 108; Indels 26; Gaps 5;
QY 6 LRRTSSSSISSTASLSVTLRRTASAPSLRRCPTSPSSSMAAASASRLVEPD 47
DB 1 MRRTVSASGLT---LSRITTSSTICWQCLRELLSNOVOIHVRKXHPTRKQVSPG 56
QY 48 AAVSASRLVEPDNQP---LINPLLVGPEMSNLTSNIRSLGSGHPSLDTVAKYVYV 104
DB 57 AAVSAQOTIFKGLPAPRPSISVDPLRIYKELKFLTKNIRQLGSGHPTLDVAKYTRRS 116
QY 105 EKGHIRPLVLMQAQTEVAPKVOGKEKVEVPVNEGSLAPPEVLNDKNDPMNMNRSGPLT 164
DB 117 EKGHMRPLVILMSQATLTPROSRNFTPSQMVNDPISSPSVLADTVPLDPLVSKSAE 176
QY 165 KQGEIRGQTSNIIASORRLAETEMHHAASLHDDVIDASETRRNAPSGNOAFGNKAL 224
DB 177 AOYDFRAGD-ENTLPQKRLAETELHTASLHDDVIDAIVRRSSANLQFGKMAVL 235
QY 225 AGDFLLGRASVALARLNEVEIELLTVIANLVGEFEMQKNTVDAIEATATQETFPDY 284
DB 236 AGDFLLGRASVALARLNEVEIELLTVIANLVGEFEMQKNTADENKPVPTDGTISY 295
QY 285 LQKTYLKTASLIKSCRASALLGATPEVADAAYAGRNIGLAFQIVDDMLDYTVSATDIL 344
DB 296 LQKTYLKTASLIKSCRAAALLGSGTPEVADAAYAGRNIGLAFQIVDDMLDYTVSGVEL 355
QY 345 GKPAGADLQIGLATAPALPAWKHAELGPMIKRKFSDGVERARELVESDGLKTRAL 404
DB 356 GKPAGADLQIGLATAPALPAWKONPGLPLVGRKFSRGGDVMAKELVYKSGVQETRAL 415
QY 405 ABEYAKALDAITFPESPARKALEQLTDKVLTRSR 440
DB 416 ABEYAKALITAVSNFPDSEAKAGLQIMCEKANRRK 451
RESULT 3
ABB83792
ID ABB83792 standard; Protein; 502 AA.
XX AC ABB83792;
XX DT 29-AUG-2002 (first entry)
XX DE Fungal decaprenyl diphosphate synthase SEQ ID NO 3.
XX KM Fungi; *Aspergillus*; *Leucosporidium*; coenzyme Q10; enzyme;
XX

decaprenyl diphosphate synthase; enzyme.
Leucosporidium scottii.
WO200252017-A1.
04-JUL-2002.
27-DEC-2001; 2001WO-JP11523.
27-DEC-2000; 2000JP-03198658.
(KANF) KANEKA CORP.
Matsuda H, Kawamukai M, Yajima K, Ikenaka Y;
WPI; 2002-500767/53.
N-PSDB; ABN81604.
Microbial production of coenzyme Q10 by transformants transferred with
fungal decaprenyl diphosphate synthase gene, on industrial scale for
application in drugs -
Claim 3; Page 33-36; 52pp; Japanese.
The invention relates to a DNA sequence (I, ABN81604-ABN81605),
containing a sequence based on that of (I) but with some bases deleted,
added, inserted and/or substituted and encoding a protein
(ABN81792-ABN83793) with decaprenyl diphosphate synthase activity or
hybridisable with (I) under stringent conditions and encoding a protein
with decaprenyl diphosphate synthase activity. The DNA is useful for the
production of coenzyme Q10 for application in drugs. The coenzyme is
simply produced on industrial scale, efficiently, with significantly
productivity.
Sequence 502 AA;
Query Match 48.5%; Score 1064; DB 23; Length 502;
Best Local Similarity 47.8%; Pred. No. 2.4e-93;
Matches 238; Conservative 80; Mismatches 112; Indels 68; Gaps 11;
QY 6 LRIRSISSIASIRSVTLRTASAPSLRLRCTPTSPSSSWAAV-----SSA 53
DB 10 LRGRAPSSLLQPTBLQKLSSTPSLRLHASPRL--SAWTAIPGLSATFPASTSTS 67
QY 54 SRLVEPDNPQPLINPLNLVGPMSNLTSIRSLGSGHPSLDTVAKYVQSEGHIRPLM 113
DB 68 SSLAGSSKVALQDPLKPLGAEMGLLRSNVQHLGSGHPALDTIAKYFFQAEKGHVRPML 127
QY 114 VLLMAQATE-VAPKVGQWEK-----VVEV-----PVNEGLAPPEVLNDKNPD 154
DB 128 ILLMSQATNGLAP---GWEQRRDQAAAAELKREGQDEGLGGDDIDEPLSPSVLNDQNS 184
QY 155 MM-----NWRSGPLTKDGELEGQTSNILASORRLAEITEMIHAASLLHDDVID 202
DB 185 MLASAKSFSDPLASLRPATPTPSIAQSIHQTHLLPSQRRLAEITEMIHAASLLHDDVID 244
QY 203 ASERRNAPSGNAGFKMMAILAGDFLLGRASVALARLNRPVEIETLATVIANLVEGEFM 262
DB 245 LAETRSAPASPSLFGNKLILAGDFLLARASLSLSLGSNEVELVASLVANLVEGEVM 304
QY 263 QLKNTV--DDAIBATA-----TQETPDYLLQKTYLKTASLAKSCRASALLGGA-- 309
DB 305 QMKGNVPGKEGLLAGAGGGSTAKGPTPEIFDHVMKTYLKTASLAKSTRATILGGCGV 364
QY 310 -----TPEVADAYAGRNGLAFQIVDDMLDYTVSATDLGRP-AGADILGLATAPAL 362
DB 365 KQWAEKGKVDKATYSVGRNLGAFQIVDDMLDFTASAAQIGKPGGAGLKLGLATAPAL 424
QY 363 FAWKHAEELGPMIKRKFSDPGDVERARELVKESDGLKTRALAEVEYAKQALDAIRTPFES 422
DB 425 YANEEFPELGAMIERKPAGEDDVEQARHLISRSRGAERTALAEHSLKARQALEGLPDS 484

QY 423 PARKALEQLTDKVLTRSR 440
DB 485 EARTALDNNMARDTLRSKK 502
RESULT 4
ID AAO14997 standard; Protein; 537 AA.
XX AAO14997;
XX 08-AUG-2002 (first entry)
XX Rhodotorula minuta decaprenyl diphosphate synthase 1.
XX Fungi; decaprenyl diphosphate synthase; enzyme; high-yield fermentation;
XX coenzyme Q10 production.
XX Rhodotorula minuta.
XX WO200240682-A1.
XX 23-MAY-2002.
XX 20-NOV-2001; 2001WO-JP10119.
XX 20-NOV-2000; 2000JP-0352940.
XX (KANF) KANEKA CORP.
XX Matsuda H, Kawamukai M, Yajima K, Ikenaka Y;
XX WPI; 2002-427097/45.
XX N-PSDB; AAL42940.
XX Decaprenyl diphosphate synthase gene of Rhodotorula origin for
XX efficient preparation of coenzyme Q10 -
XX Claim 3; Page 33-36; 50pp; Japanese.
XX The invention comprises the amino acid and coding sequences of fungal
XX (Rhodotorula minuta) decaprenyl diphosphate synthase enzymes. The fungal
XX decaprenyl diphosphate synthase DNA and protein sequences are useful for
XX the efficient, high-yield fermentative production of coenzyme Q10 - for
XX pharmaceutical use. The present amino acid sequence represents a
XX Rhodotorula minuta decaprenyl diphosphate synthase enzyme.
XX Sequence 537 AA;
Query Match 46.9%; Score 1028.5; DB 23; Length 537;
Best Local Similarity 48.6%; Pred. No. 6.9e-90;
Matches 237; Conservative 84; Mismatches 108; Indels 59; Gaps 12;
QY 2 ASPALIRSISSRSIASLSRVT-LRTASAPS-----LRLRCTPTSPSSSWAA 48
DB 60 SKPAARSTSTAPALPSPSTSDPQSSSPSSSSSSSILPDLFRPLSSSSSSSTSS 119
QY 49 AVSASRLVEPDNPQPLINPLNLVGPMSNLTSIRSLGSGHPSLDTVAKYVQSEGHK 108
DB 120 SSSSSSNRKNNTSNITFDPLQVLGNELSLRSNVQALLGSGHPALDTIAKYFFQAEKGK 179
QY 109 IRPLMYLLMAQATE-VAPKVGQWEKVEV-----PVNEGL--APPEVLNDKNP 153
DB 180 IRPMIVLLMSQATNGLAP---GFEERSKLELSGRKQTDPSRINDPLEVKADILNDSNP 236
QY 154 -DMNWRSGPLTKDGELEGQTSNILASORRLAEITEMIHAASLLHDDVIDASERRNAPS 212
DB 237 SSFAASSSSPL-----DSMFSTSNVLPFSQRRLAEITEMIHAASLLHDDVIDGSAMRRAQS 292
QY 213 GNOAFGNKMAILAGDFLLGRASVALARLNRPVEIETLATVIANLVEGEFMOLKNTVDDAI 272
DB 293 APAAFGNKISVLGGDFLLARASLYLSRLGSGNEVELVASLVANLVEGEVMQIKG---NAP 349

Qy	273	ETA-----TOETFDYVLOKTYKTSIAKSRASALIGATPE-----V	313
Db	350	ESNAGSGSEVAVHRLTPELFEHYMKTYIKTSLINKSRATTTIIGAGCKGWIIGERI	409
Qy	314	ADAAVAYGENTLGAFOIVDDMLDYVSATDLGKPA-GADLQGLATAPALFAWKHAETG	372
Db	410	KDIASVYSGNLTGIAFQLVDDLLDFTATDAQFGKPSQADLKLGLATAPALVAMEEFPENG	469
Qy	373	PMIKKFSFGPEVERARELVEKSDGLEKRALAEVQAOLAIRTPESPARKALEQUT	432
Db	470	QMLIKKFEFEGVEETPARLVRSAGEKTVKAEKPAALAMELOGLPSPDAREALEGLT	529
Qy	433	DKVLTCSR 440	
Db	530	KTVLNRTK 537	
RESULT 5			
ID	AA014998	standard; Protein; 425 AA.	
AC	AA014998;		
DT	08-AUG-2002	(first entry)	
DE	Rhodotorula minuta decaprenyl diphosphate synthase 2.		
KW	Fungi; decaprenyl diphosphate synthase; enzyme; high-yield fermentation;		
KM	coenzyme Q10 production.		
OS	Rhodotorula minuta.		
PN	WO200240682-A1.		
PD	23-MAY-2002.		
PF	20-NOV-2001; 2001WO-JP10119.		
PR	20-NOV-2000; 2000JP-0352940.		
PA	(KANF) KANEKA CORP.		
PI	Matsuda H, Kawamukai M, Yajima K, Ikenaka Y;		
DR	WPI; 2002-427097/45.		
DR	N-PSDB; AAL42941.		
PT	Decaprenyl diphosphate synthase gene of Rhodotorula origin for		
PT	efficient preparation of coenzyme Q10		
PS	Claim 4; Page 39-41; 50pp; Japanese.		
XX	The invention comprises the amino acid and coding sequences of fungal		
XX	(Rhodotorula minuta) decaprenyl diphosphate synthase enzymes. The fungal		
XX	decaprenyl diphosphate synthase DNA and protein sequences are useful for		
XX	the efficient, high-yield fermentative production of coenzyme Q10 for		
XX	pharmaceutical use. The present amino acid sequence represents a		
XX	Rhodotorula minuta decaprenyl diphosphate synthase enzyme.		
SQ	Sequence	425 AA;	
Query Match 45.5%; Score 999; DB 23; Length 425;			
Best Local Similarity 52.9%; Pred. No. 3.3e-87;			
Matches 221; Conservative 66; Mismatches 85; Indels 46; Gaps 10;			
Qy	59	PDPNOLINPLNLVGPENGLTNTSLIGSGHPSLDYAKYVSGEGHIRPLMLMA	118
Db	18	PNALNTIFPDLPLVGNELSLRSNVQALLGSGHPALDTIAKYFQAEGHIRPMIVLMS	77
Qy	119	GATE-VAPKVGMEKVVVE-----PVNEG--APPELVNDKRP-DMNNRSGP	162
Db	78	QATNGIAP---GFEESKSLSEGRKOTDPSRSINDPLEVKADEIINDSPSFAASSSP	134

Oy	163	LTKGEIEGQCSNIIASQRRLAEITEMIHAASLHDDVIDAETRRNAPSQOAFGNMA	222
Db	135	L-----DSMPSTSNVLPSGRRLAEITEMIHVASLSLHDVDIDSAMRRAQASAPAAKNNIS	190
Oy	223	ILAGDFLLGRASVALARLNRPVEILLATVIANLVESGFMOIKNTVDALIENTA-----	276
Db	191	VLGGLFLLARSLLYISRSGSNEVELVASVANIVEGVMOIKG--NAPESNAGSKEV	247
Oy	277	-----TOSTPPLYOKTVYLTASTLIASKCRASALLGGATPE-----VADAAAYAGR	323
Db	248	AVHRLTFEIFEHYMKTYLTAKSLIASTRATTILGGAGEKQGI EGERIKDIAYSGRN	307
Oy	324	LGAFOIVDDMLDYVSATDLEGKPA-GADLOGLATAPALFAWKHAELGPWIKKFSDP	382
Db	308	LGIATQLVDDLDFPATDAQFGPKSQGADKLGLATAPALVAMEEPFMGMILRKFE	367
Oy	383	GDEVARLARLVESDGLKTRRALAEYAQAQALDAIRTPESPARKALEQTDKYLTRSR	440
Db	368	GDVEATARNLVKSAGPEKTIVLAEKHALAMBALGGLPESDAREALFGLTKIVLNRTX	425
 RESULT 6 AA014999 ID AA014999 standard; Protein; 403 AA. AC XX AC AA014999; XX DT 08-AUG-2002 (first entry) XX DE Rhodotorula minuta decaprenyl diphosphate synthase 3. XX KW Fungi; decaprenyl diphosphate synthase; enzyme; high-yield fermentation; XX coenzyme Q10 production. XX OS Rhodotorula minuta. XX EN WO200240682-A1. XX PD 23-MAY-2002. XX PE 20-NOV-2001; 2001WO-JP10119. XX PR 20-NOV-2000; 2000JP-0352940. XX PA (KANF) KANEKA CORP. XX PI Matsuda H, Kawamukai M, Yajima K, Ikenaka Y; XX WPI; 2002-427097/45. DR N-PDSB; AAL42942. XX PT Decaprenyl diphosphate synthase gene of Rhodotorula origin for XX efficient preparation of coenzyme Q10 - XX PS Claim 19; Page 44-46; 50pp; Japanese. XX CC The invention comprises the amino acid and coding sequences of fungal CC (Rhodotorula minuta) decaprenyl diphosphate synthase enzymes. The fungal CC decaprenyl diphosphate synthase DNA and protein sequences are useful for CC the efficient, high-yield fermentative production of coenzyme Q10 - for CC pharmaceutical use. The present amino acid sequence represents a CC Rhodotorula minuta decaprenyl diphosphate synthase enzyme. XX SQ Sequence 403 AA;			
Oy	65	LINPLNLVGPEKSNLTSNRISLIGSGHPSLDTPVAKYYVQSBEKHRIPLMVLMAQTE-V	123
Db	2	IIFDPLQVLGNELSSRSNVQALSGCHALDITAIKYVOABEKHIRPMIVLMSGATNCL	61
 Query Match 45.4%; Score 996; DB 23; Length 403; Best Local Similarity 53.4%; Pred. No. 5,9e-87; Matches 220; Conservative 65; Mismatches 81; Indels 46; Gaps 10.			

XX
PN
XX

WO200063391-A2

XX AAU72775;
 XX 26-FEB-2002 (first entry)
 XX Arabidopsis tocopherol cyclase ATP8.
 DE
 XX Tocopherol cyclase; isoprenoid; carotant; cytosolic; noctropic;
 KM neuroprotective; ophthalmological; antiarthritic; anti-aging;
 KM antioxidant; food additive; seed oil; tocotrienol; free radical damage;
 KM cardiac diseases; cancer; cataract; retinopathy; Alzheimer's disease;
 KM neurodegeneration; arthritis; transgenic plant.
 XX Arabidopsis sp.
 OS
 XX MO200179472-A2.
 PN
 XX 25-OCT-2001.
 PD
 XX 13-APR-2001; 2001WO-US12334.
 PF
 XX 14-APR-2000; 2000US-0549848.
 PR 14-OCT-2000; 2000US-0668069.
 XX (MONS) MONSANTO TECHNOLOGY LLP.
 PA
 XX Subramaniam SS, Slater SC, Karberg K, Chen R, Valentin HE;
 PI Wong YH;
 PT N-PSDB; AAS96924.
 DR
 XX WPI; 2002-066365/09.
 DR N-PSDB; AAS96924.
 PT New nucleic acids encoding tocopherol cyclase, useful in increasing the
 production or modulation of tocopherol cyclase in a host plant,
 particularly those plants that produce vegetable oils for edible and
 industrial uses -
 PT
 XX Example 1; Page 113-114; 166pp; English.
 PS
 XX The invention relates to an isolated nucleic acid sequence encoding a
 tocopherol cyclase, an enzyme involved in the synthesis of
 CC tocopherol, an antioxidant isoprenoid. Also included are transgenic
 CC plants expressing tocopherol cyclase where the tocopherol is purified
 CC from the seed as a refined and deodorized oil produced by treating a
 CC tocopherol rich oil by distilling under low pressure and high
 CC temperature, where the refined oil has reduced free fatty acids and a
 CC substantial percentage of tocopherol present in the pretreated oil.
 CC The nucleic acid is useful in the increased production or
 CC modulation of tocopherol cyclase in a host plant or plant cell. The
 CC polynucleotides can be used as hybridisation probe for RNA, cDNA or
 CC genomic DNA to isolate full length DNAs or genomic clones encoding a
 CC polypeptide and to isolate cDNA or genomic clones of other genes having
 CC high sequence identity to the given polynucleotide sequences.
 CC Tocopherols and tocotrienols are useful as antioxidants and play an
 CC important role in protecting cells from free radical damage, in the
 CC prevention of cardiac diseases, cancer, cataract, retinopathy,
 CC Alzheimer's disease, neurodegeneration, and arthritis, and in
 CC anti-aging. The nucleic acid constructs are useful for the increased or
 CC decreased expression of tocopherol cyclase in plants involved in the
 CC production of vegetable oils for edible and industrial uses. The
 CC present sequence represents a tocopherol cyclase or tocopherol cyclase
 CC -like protein.
 CC
 XX
 SO Sequence 321 AA;

Query Match 30.0%; Score 658; DB 23; Length 321;
 Best Local Similarity 42.0%; Pred. No. 1.5e-54;
 Matches 150; Conservative 64; Mismatches 93; Indels 50; Gaps 5;

QY 92 PSIDTVAKYUUS--EGNHRPLWVLLMAQATEVAPKVGQEKVVEVNVNGLAPRVLN 149
 DB 7 PKLASAAEFPPKRGVQKQFSTILLMAATLVN-----RVPPALI 47

QY 150 DKNPDMNMRSGPLTKDGEIEGOTSIIASQRIAEITTEMIHAASLHDVDVDASETRN 209
 DB 48 GSTDIV-----TSELVRQGIABIEITTEMIHVASLHDVDVDDADTBRG 91
 QY 210 APGNOAFGNKAILAGDFELGRASVALARLNPEVIELATVIANLVEGEMQKNTVD 269
 DB 92 VGSLLVNMGNKNSVLAGDFELSRACALALAKNTEVALLATRAVEHLVIGEMETLS--- 148
 QY 270 DAIEATATGETPDYVYOKTYLTKASTIAKSCASALLGGATPEVADAAYAGRNGLAFQ 329
 DB 149 ---STEGRYSMDYVWQKTYVYASTIASNSCAVAVLNGQTAEVAVLAFEGYRNGLAFQ 204
 QY 330 IYDMDLDTYVSNATDCKPRAGADLQGLATAPALPMKHAELGPMIKKFPSPGVPEAR 389
 DB 205 LIDDLIDFTGTSASLGKSLDIRHGVITAPLTFAMEEPPQIREVVDVEKPRVDIAL 264
 QY 390 ELVEKSDGLEKTRALAEYQAQALDAIRFPESP-----ARKALEQITDKVLTSSR 440
 DB 265 EYIGKSKGIQRARELAMEHANLAAAIGSLPETDNEVDYKRSRALIDLTRVITNKK 321

RESULT 13
 ABP41763
 ID ABP41763 strandcd; Protein; 272 AA.
 XX
 AC ABP41763;
 XX
 DT 22-AUG-2002 (first entry)
 XX
 DE Human ovarian antigen HHNG108, SEQ ID NO:2895.
 XX
 KM Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KM ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KM infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KM PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KM inflammatory condition; immune disorder; blood disorder;
 KM cardiovascular disorder; respiratory disorder; neurological disorder;
 KM gastrointestinal disorder; urinary system disorder; drug screening;
 KM gene therapy; chromosome mapping; forensic analysis;
 KM antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KM antiinflammatory; gynaecological; reproductive; chromosome 10p11.2-12.
 XX
 OS Homo sapiens.
 XX
 PN WO200200677-A1.
 PN
 PD 03-JAN-2002.
 PD
 PF 07-JUN-2001; 2001WO-US18569.
 PF
 XX 07-JUN-2000; 2000US-209467P.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Birse CE, Rosen CA;
 PI WPI; 2002-147878/19.
 DR N-PSDB; ABQ54840.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 PT
 XX
 PS Claim 11; SEQ ID NO 2895; 2922bp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC recombinant polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,

PN WO200233060-A2.
 XX
 PD 25-APR-2002.
 XX
 PF 12-OCT-2001; 2001WO-US42673.
 XX
 PR 14-OCT-2000; 2000US-0668071.
 XX
 PA (MONSANTO TECHNOLOGY LLC.
 XX
 PI Lasserer MM, Savidge B, Weiss JD, Mitsky TA, Post-Beltemmiller MA;
 PI Valentin HE;
 XX
 DR WPI; 2002-463312/49.
 XX
 PT Novel nucleic acid sequences encoding prenyltransferase derived from
 PT eukaryotic and prokaryotic sources useful for producing plants and
 PT seeds with altered tocopherol content and compositions -
 XX
 PS Claim 6; Page 110-111; 148pp; English.
 XX
 CC The invention relates to a novel isolated nucleic acid sequence encoding
 CC a prenyltransferase (PTS). The nucleic acid construct of the invention is
 CC useful for altering isoprenoid content and for producing an isoprenoid
 CC compound of interest in a host cell, such as *Synechocystis* sp., or a
 CC plant cell obtained from *Arabidopsis*, corn, rice, wheat, leek, canola,
 CC cotton or tomato. The construct is also useful for increasing
 CC biosynthetic flux in a host cell toward production of an isoprenoid
 CC compound such as tocopherol and tocotrienol. The polynucleotide is useful
 CC for producing plants and plant parts, such as seeds, enriched in
 CC tocopherols. The sequence shown represents *Arabidopsis* sp. ATP1, a
 CC polytransferase identified in the invention.
 XX
 SQ Sequence 297 AA;
 Query Match 25.6%; Score 561; DB 23; Length 297;
 Best Local Similarity 37.5%; Pred. No. 2.8e-45;
 Matches 134; Conservative 61; Mismatches 88; Indels 74; Gaps 6;
 QY 92 PSIDTVAKYVQ--EGKHIRPLNVLMAQTEVAPKVGWEKVEVNVNEGAPPEVLN 149
 Db 7 PKLASAEYFFKRGVQKQFPRSTILLMATLVN-----RVPEALI 47
 QY 150 DKNPDMNMMSGFLTKDGELEGTSNLLASQRLAEITTEMIHAASLLHDVYIDASETRRN 209
 Db 48 GESTDIV-----TSELRVQRGIAEITEMIHVASLHDDVDADTRRG 91
 QY 210 APSGNQAFGNMALLAGDFLLGRASVALARLNPEVIELLATVIANLVEGPFMQLKNTVD 269
 Db 92 VGSINVMGNK-----VVALLATVAEHLVTGETMEITS--- 124
 QY 270 DATEATATQETFDYVLYLQTYLTKTASLIKSCRASALGATPEVADAAYAGRNIGLAFQ 329
 Db 125 -----STFORYSMDYTKYTKTASLISNSCKAVAVLTGQTAEVAVLAFEGRNIGLAFQ 180
 QY 330 IVDMDLYVSATDLGKPGADLQGLATAPALPAMKHAELGPMIKKFSDDGVDERAR 389
 Db 181 LIDDIPLFTGTSASLGKGSLSDIRHGVTAPILPAMEEFPQLRVVDQVEKDPNNVDIAL 240
 QY 390 ELVKSQGLEKTRALAEYAKALDAIRTPESP-----ARKALEQLTDKVLTRSR 440
 Db 241 EYLGKSKGIQARELAMEHANLAAALGSLPETNEDVKSRRLIDLTHRVITRNK 297

Search completed: January 16, 2003, 09:46:48
 Job time : 58 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:46:36 ; Search time 24 Seconds
(without alignments)
539.420 Million cell updates/sec

Title: US-09-830-111A-2
Perfect score: 2194
Sequence: 1 MASPALRIRSISSRIASLR.....ESPAKALEQLTKVLTISR 440

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCUS COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	445.5	20.3	333	4	US-09-025-819-29
2	445.5	20.3	333	4	US-09-808-126-29
3	445.5	20.3	333	4	US-09-803-951-29
4	439.5	20.0	335	4	US-09-673-018-2
5	409	18.6	320	3	US-09-217-609A-6
6	409	18.6	320	4	US-08-873-235B-6
7	397.5	18.1	325	3	US-09-009-895-2
8	346	15.8	357	4	US-09-134-001C-3250
9	336.5	15.3	325	3	US-09-217-609A-2
10	336.5	15.3	325	4	US-08-873-235B-2
11	283	12.9	393	3	US-09-187-050-28
12	280	12.8	393	3	US-09-187-050-27
13	278	12.7	393	3	US-09-187-050-2
14	278	12.7	393	3	US-09-187-050-14
15	278	12.7	393	3	US-09-187-050-16
16	278	12.7	393	3	US-09-187-050-18
17	278	12.7	393	3	US-09-187-050-20
18	278	12.7	393	3	US-09-187-050-22
19	278	12.7	393	3	US-09-187-050-24
20	278	12.7	393	3	US-09-187-050-26
21	278	12.7	393	3	US-09-187-050-29
22	278	12.7	393	3	US-09-187-050-30
23	278	12.7	393	3	US-09-187-050-31
24	278	12.7	393	3	US-09-187-050-32
25	278	12.7	393	3	US-09-187-050-33
26	274	12.5	393	3	US-09-187-050-34
27	264	12.0	285	3	US-09-187-050-12

28	260	11.9	330	1	US-08-410-167A-4	Sequence 4, Appli
29	260	11.9	330	2	US-08-898-560-1	Sequence 1, Appli
30	258	11.8	330	4	US-09-101-126-1	Sequence 1, Appli
31	252	11.5	377	4	US-09-420-211-2	Sequence 2, Appli
32	248	11.3	297	1	US-08-534-910B-7	Sequence 7, Appli
33	248	11.3	297	1	US-08-534-910B-8	Sequence 8, Appli
34	248	11.3	297	1	US-08-534-910B-10	Sequence 10, Appli
35	247	11.3	297	1	US-08-534-910B-6	Sequence 6, Appli
36	247	11.3	297	3	US-08-886-466-2	Sequence 2, Appli
37	247	11.3	297	4	US-09-475-304-2	Sequence 3, Appli
38	247	11.3	297	4	US-09-101-126-3	Sequence 3, Appli
39	247	11.3	297	4	US-09-367-528A-5	Sequence 5, Appli
40	243	11.1	297	4	US-09-367-528A-1	Sequence 1, Appli
41	241	11.0	297	4	US-09-367-528A-3	Sequence 3, Appli
42	239	10.9	291	4	US-09-275-742-2	Sequence 2, Appli
43	238	10.8	297	1	US-08-534-910B-9	Sequence 9, Appli
44	206	9.4	298	1	US-08-095-726-4	Sequence 4, Appli
45	206	9.4	298	1	US-08-096-043-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-025-819-29
; Sequence 29, Application US/09025819
; Patent No. 6225097
; GENERAL INFORMATION:
; APPLICANT: Obata, Shusei
; APPLICANT: Nishino, Tokuzo
; APPLICANT: Koyama, Tanetoshi
; APPLICANT: Sato, Yoshihiro
; TITLE OF INVENTION: DECAPRENYL DIPHOSPHATE SYNTHETASE GENE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KENYON & KENYON
; STREET: 1500 K Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/025,819
FILING DATE: 19-FEB-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 251675
FILING DATE: 17-SEP-1997
ATTORNEY/AGENT INFORMATION:
NAME: Khalilian, Hourl
REGISTRATION NUMBER: 39,546
REFERENCE/DOCKET NUMBER: 10235/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-220-4200
TELEFAX: 202-220-4201
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-025-819-29

Query Match 20.3%; Score 445.5; DB 4; Length 333;
Best Local Similarity 30.3%; Pred. No. 4.5e-35;
Matches 120; Conservative 64; Mismatches 125; Indels 87; Gaps 9;
QY 62 NQPLINPLNVGPE---MSNLTSNIRSLIGSGH-PSLDTVAKYVYQSECKHIRPLMVL 116

```

Db      4 NENSVKPLDRISVELAGMDRVNALLIRRMASRHAPRIPEVTAHLVEAGGKRLPMLVL- 62
QY      117 MAQATEVAPKVOGWEKVEVFNVEGLAPPEVLNDKNDPMNMNRSGPLTKDGEIEGQTSNI 176
Db      63 -----AAARLCGYQ-----GNSHV 76
QY      177 LASQRRLAETITMTHAASLHDDVIDASSTRRNAPSGNOAFGNKMAILLAGDFLLGRASVA 236
Db      77 L-----LAAAVEFIHTATLTHDDVDDESQGRGRTANLMDNKS SVLVGDYLFARSFOL 131
QY      237 LARLRNPEVIELLATVIANLVEGEFMOLKNYVDDAIEATQOE--TFDYLOKTYLTKTA 293
Db      132 MADTESQVWRILANASATIAEGEVLOL-----TAAQDVSTTEDTYIQIVRGKTA 181
QY      294 SLIAKSCRASALLGATPEVADAAYAGRNIGLAFQIVDDMLDYVSATDLGKPGADLO 353
Db      182 ALFSAATEAGAVAGADPAVQCALFDYGDALGIAFQIVDDLLDYGSGTTTIGKVGDDFR 241
QY      354 LGIATAPALFA-----WKHHAELGPMIKKFSDPGDERARELVEKSDGLEKTR 402
Db      242 ERKLTLPVIAKAIARADEARAFMER-----TIGQGRQDEADLATALEILRRREALLEAAR 295
QY      403 ALAEVYAKALDAIRTFESPARKALEQLTDKVLTR 438
Db      296 ADAIAMAGRAKALQAPDQPLRILADLADPVS 331

```

RESULT 2

```

US-09-808-126-29
Sequence 29, Application US/09808126
Patent No. 6410280

```

GENERAL INFORMATION:

APPLICANT: Obata, Shusei

Nishino, Tokuzo

Koyama, Tanetoshi

Sato, Yoshihiro

TITLE OF INVENTION: DECAPRENYL DIPHOSPHATE SYNTHETASE GENE

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: KENYON & KENYON

STREET: 1500 K Street, N.W.

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/808,126

FILING DATE: 08-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/025,819

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Khalilian, Hourti

REGISTRATION NUMBER: 39,546

REFERENCE/DOCKET NUMBER: 10235/2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-220-4200

TELEFAX: 202-220-4201

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:

LENGTH: 333 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 29:

```

Query Match      20.3%; Score 445.5; DB 4; Length 333;
Best Local Similarity 30.3%; Pred. No. 4.5e-35;
Matches 120; Conservative 64; Mismatches 125; Indels 87; Gaps 9;
QY      62 NOPLINPLNVGPE-----MSNLTNSIRSLGSGH-PSLDYAKYVYQSEGKHIRPLMVL 116
Db      4 NENSVKPLDRISVELAGMDRVNALLIRRMASRHAPRIPEVTAHLVEAGGKRLPMLVL- 62
QY      117 MAQATEVAPKVOGWEKVEVFNVEGLAPPEVLNDKNDPMNMNRSGPLTKDGEIEGQTSNI 176
Db      63 -----AAARLCGYQ-----GNSHV 76
QY      177 LASQRRLAETITMTHAASLHDDVIDASSTRRNAPSGNOAFGNKMAILLAGDFLLGRASVA 236
Db      77 L-----LAAAVEFIHTATLTHDDVDDESQGRGRTANLMDNKS SVLVGDYLFARSFOL 131
QY      237 LARLRNPEVIELLATVIANLVEGEFMOLKNYVDDAIEATQOE--TFDYLOKTYLTKTA 293
Db      132 MADTESQVWRILANASATIAEGEVLOL-----TAAQDVSTTEDTYIQIVRGKTA 181
QY      294 SLIAKSCRASALLGATPEVADAAYAGRNIGLAFQIVDDMLDYVSATDLGKPGADLO 353
Db      182 ALFSAATEAGAVAGADPAVQCALFDYGDALGIAFQIVDDLLDYGSGTTTIGKVGDDFR 241
QY      354 LGIATAPALFA-----WKHHAELGPMIKKFSDPGDERARELVEKSDGLEKTR 402
Db      242 ERKLTLPVIAKAIARADEARAFMER-----TIGQGRQDEADLATALEILRRREALLEAAR 295
QY      403 ALAEVYAKALDAIRTFESPARKALEQLTDKVLTR 438
Db      296 ADAIAMAGRAKALQAPDQPLRILADLADPVS 331

```

RESULT 3

```

US-09-803-951-29
Sequence 29, Application US/09803951
Patent No. 6413761

```

GENERAL INFORMATION:

APPLICANT: Obata, Shusei

Nishino, Tokuzo

Koyama, Tanetoshi

Sato, Yoshihiro

TITLE OF INVENTION: DECAPRENYL DIPHOSPHATE SYNTHETASE GENE

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: KENYON & KENYON

STREET: 1500 K Street, N.W.

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/803,951

FILING DATE: 13-Mar-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/025,819

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Khalilian, Hourti

REGISTRATION NUMBER: 39,546

REFERENCE/DOCKET NUMBER: 10235/2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-220-4200

TELEFAX: 202-220-4201

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:

LENGTH: 333 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-803-951-29

Query Match 20.3%; Score 445.5; DB 4; Length 333;
Best Local Similarity 30.3%; Pred. No. 4.5e-35;
Matches 120; Conservative 64; Mismatches 125; Indels 87; Gaps 9;
QY 62 NOPLNPLNLVGPB---MSNLSTNIRSLGSGH-PSLDTVAKYVQSECKHIRELMVLL 116
DB 4 NENVKPLDLRLSVELAGDMRVNALIRERWASHPRIPEVTAHLVZAGGRRLRPLMLVL- 62
QY 117 MAQATEVAPKVOGMEKVVPEVNEGLAPPEVLNDKNPDMNMRSGPLTKDGEIEGQTSNI 176
DB 63 -----AAARLCGVQ-----GNSHV 76
QY 177 LASORRLAEITEMIHAASILLHDDVIDASSTRNAPSNGQAFGNKMAILAGDFLLGRASVA 236
DB 77 L-----LAAAVEFIHTATLLHDDVDVDESQORRGRTPTANLLWLNKSSVLVGDYLFARSQ 131
QY 237 LARLRNPEVIELLATVIANIVEGEFMOLKNTVDDAIEATATQE---TFDYVLOKTYLKTA 293
DB 132 MATDESQVNRILANASATTAEVQLQ-----TAAQDVSTTETDYIQIVRGKTA 181
QY 294 SLIAKSCRASLAGGATPEVADAAAYAGRNLGLAFQIVDDMLDYTVSATDLGKPGAGADLQ 353
DB 182 ALFSAATEAGAVAGADPAVQQAQFDYGDALGIAFQIVDDLLDYGGSTTTIGKNGVDDFR 241
QY 354 LGLATAPALFA-----WKHHAELGPMIKRKFSDPGDVERARELVEKSDGLEKTR 402
DB 242 ERKLTLPVKAIRADEAFERFWR-----TIGQGRQDEADLATALEILRRREALEAAR 295
QY 403 ALAEVYAKALDAIRTPESPARKALEQTLTKVLTR 438
DB 296 ADAIAWAGRAKALQAAPDQPLRRILADLADVVS 331

RESULT 4
US-09-673-018-2
; Sequence 2, Application US/09673018
; Patent No. 6461842
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Hideyuki
; APPLICANT: KAWAMUKAI, Makoto
; APPLICANT: YAJIMA, Kazuyoshi
; APPLICANT: IKENAKA, Yasuhiro
; APPLICANT: NISHI, Kenichi
; APPLICANT: HASEGAWA, Junzo
; APPLICANT: TAKAHASHI, Satomi
; TITLE OF INVENTION: METHOD FOR PREPARING COENZYME Q10
; FILE REFERENCE: 2000-1379A/LC/00177
; CURRENT APPLICATION NUMBER: US/09/673, 018
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: JP 11/32657
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Agrobacterium sp.
US-09-673-018-2

Query Match 20.0%; Score 439.5; DB 4; Length 335;
Best Local Similarity 30.3%; Pred. No. 1.7e-34;
Matches 115; Conservative 62; Mismatches 140; Indels 65; Gaps 5;
QY 63 QPLINPL-NLVGPMNSLNTNIRSLGSGHPSLDTVAKYVQSECKHIRELMVLLMQAT 121
DB 11 QASVKPLVDLTSPDMERVNQLILSRAGSDVQMPEVANHLISSGGKRLRPMLTLASA--- 67

QY 122 EVAPKVOGMEKVVPEVNEGLAPPEVLNDKNPDMNMRSGPLTKDGEIEGQTSNIIASQR 181
DB 68 ----AMFGYEGDAHV----- 78
QY 182 RLAEITEMIHAASILLHDDVIDASSTRNAPSNGQAFGNKMAILAGDFLLGRASVALARLR 241
DB 79 KLATSVFPHWTATLLHDDVDVDESLLRGGKSTARTWGNQASVLVGDYFLGQAFRMMVDVG 138
QY 242 NPEVIELLATVIANIVEGEFMOLKNTVDDAIEATATQETFDYVLOKTYLKTAASLIASCR 301
DB 139 SLDALDVLSTAASVIAEGEVQLS-----VAKNMETTEDDYLVIRAKTAALFAAAAE 191
QY 302 ASALLGGATPEVADAAAYAGRNLGLAFQIVDDMLDYTVSATDLGKPGAGADLQGLATAPA 361
DB 192 VGPVIAQTAKSDRNALKSYGMNLGLAFQLVDDVLDYGGASDLGKNGVDDFREGKITLPV 251
QY 362 LFAWKHHAELGPMIKRKFSDPGD-----VERARELVEKSDGLEKTRALAEVYAKALDAI 416
DB 252 ILSVREGTPEDFRWFREAEIEGGSDSNLRLALGLIKRYGGLSDTTIARAQHYGTIARDAL 311
QY 417 RTPESPARKALEQTLTKVLTR 438
DB 312 APLPVPSPKSALEVEVIDFCIDR 333

RESULT 5
US-09-217-609A-6
; Sequence 6, Application US/09217609A
; Patent No. 6071733
; GENERAL INFORMATION:
; APPLICANT: MURAMATSU, Masayoshi
; APPLICANT: KOIKE, Ayumi
; APPLICANT: OGURA, Kyoza
; APPLICANT: KOYAMA, Tanetoshi
; APPLICANT: SHIMIZU, Naoto
; APPLICANT: CHO, Yenmin
; TITLE OF INVENTION: Prenyl Diphosphate Synthetase Genes
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, NW - Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
; SOFTWARE: WordPerfect 6.1 for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/217,609A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/873,235
; FILING DATE: 11-Jun-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: TOFFENETTI, Judith L.
; REGISTRATION NUMBER: 39,048
; REFERENCE/DOCKET NUMBER: 10235/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-429-1776
; TELEFAX: 202-429-0796
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-217-609A-6

Qy	70	NLVGEMSNLTSNRKSLGSHPSLDIVAKYVQSEGHRIPELWLLMAQATEVAPKVG	129
Db	8	SFLSDLLAAEEELERAAVQSEYEPGLGMAHLTLQAGCKRIIPVFLLAARF	58
Qy	130	MEKVVEPVNENGLAPPEVLNKNPDMMNRSGPLTKGKEIGQCSNLLAQBRRLAETEM	189
Db	59	-----GQTD-----LEKMKHVAVALTEL	75
Qy	190	IHAASLHDVDVIDASETRRNAPSGNOAFGNKMAILAGDFLGRASVALARLRNPEVIELL	249
Db	76	IHMASLVHDDVIDADADLERGRPTIKAKMSNRFPAYTGDVLEFARSLERMAELGNPRAHQVL	135
Qy	250	ATVLANLVGEFMQAKTVDALEATATQEFEDY----YLQKTYLTKASLLAASCRASL	305
Db	136	AKTIVEVGRGEIQRD-----KYRFDPFLRTYLRIRRKXALLIASSCOQAL	184
Qy	306	LGAATPEVADAAYAAGRNGLAFQIYDDMLDYTSATDGLKPGADATQLGIATAPALFAM	365
Db	185	AAGAEPIYKRLVYMGHVGKMSFOITDDIIDFTGTEQGLKPGASDILLQGNVLLPVLYAL	244
Qy	366	KHH-----AELGPMIKKKSFDPGDVEFARLVEKSDGLEKTRALAEEYQAKALDIRTF	419
Db	245	SDERVKAAIIAIVGE-----TDVAEMAVVISAIRKTDALIERSYALLSDRYLDKALLHLDGL	299
Qy	420	PESPARKALEQL	431
Db	300	PMNEARKGLLRDI	311

RESULT 6
 US-08-873-235B-6
 Sequence 6, Application US/08873235B
 Patent No. 6174715
 GENERAL INFORMATION:
 APPLICANT: MORAMATSU, Masayoshi
 APPLICANT: KOIKE, Ayumi
 APPLICANT: OGURA, Kyoze
 APPLICANT: KOYAMA, Tanetoshi
 APPLICANT: SHIMITSU, Naoto
 APPLICANT: CHO, Yewlin
 TITLE OF INVENTION: Prenyl Diphosphate Synthetase Genes
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Kenyon & Kenyon
 STREET: 1025 Connecticut Avenue, NW - Suite 600
 CITY: Washington
 STATE: DC
 COUNTRY: US
 ZIP: 20036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
 SOFTWARE: Wordperfect 6.1 for Windows
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/873,235B
 FILING DATE: 11-Jun-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 154441/1996
 FILING DATE: 14-Jun-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: TOFENETTI, Judith L.
 REGISTRATION NUMBER: 39,048
 REFERENCE/DOCKET NUMBER: 10235/1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-429-1776
 TELEFAX: 202-429-0796
 INFORMATION FOR SEQ ID NO: 6:

```

; SEQUENCE CHARACTERISTICS:
;     LENGTH: 320 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
;
US-08-873-235B-6

```

[illegible]

```

RESULT 7
US-09-009-895-2
; Sequence 2, Application US/09009895
; Patent No. 6103488
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Hideyuki
; APPLICANT: KAWAMUKAI, Makoto
; APPLICANT: NAKAGAWA, Tsuyoshi
; APPLICANT: TANAKA, Tatsunori
; TITLE OF INVENTION: METHOD OF FORMING UBIQUITONE-10
; FILE REFERENCE: Matsuda9
; CURRENT APPLICATION NUMBER: US/09/009,895
; CURRENT FILING DATE: 1998-01-21
; EARLIER APPLICATION NUMBER: 09-244811
; EARLIER FILING DATE: 1997-08-27
; NUMBER OF SEQ. ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Rhodobacter capsulatus
US-09-009-895-2

Query Match      18.1%; Score 397.5; DB 3; Length 325;
Best Local Similarity 30.4%; Pred. No. 2e-30;
Matches 116; Conservative 59; Mismatches 142; Indels 65; Gaps 8;

QY      60  DPNOPLINPLNVGPENSNLTNSIRSLGSGHPISDITVAKYVYSGGKHRIPLMVLMAQ 119
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      4  DFKODILAP---VAQDFPAMDQFINEGISKVALVMSYSKSHVEAGGRMBPIMCLLAAY 60
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      120 ATEVAPKYGKGEKVEYEVFVNEGFLAPPEVYLNDKNPDMNMRSGPLTKDGEISGQSNILAS 179
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

Db 61 AC-----GETN--LKH 69

QY 180 ORLAITEMIHAASLLHDDVIDASSTRNAPSGNQAFGNKMAILAGDFLLGRASVALAR 239

Db 70 AOKLAIIEMLTATLVHDDVDDESGLRRGPTANATWNNQTAVALVGVDFLIARAFOLLVD 129

QY 240 LRNPEVELLATVIANIVEGEFMOLKNTVDDAIEATATQETFDYLOKTYLKTASLIAS 299

Db 130 LDNMILKDFSTGCEIAEVEIQLQ-----AQHQPPTTE--DIYLOIHHGKTSRUFELA 182

QY 300 CRASALLGGATPEVADAAAYGRNGLIAGFOIVDDMLDVTVSATDLGKPGAGADLQGLATA 359

Db 183 TEGAAAILAG-KPEVREPLRFAGFGNAFOIIDDILDTSDATLGLKNGICDDLMGKPTL 241

QY 360 PALFANKH-HAELGPMIKRKFSPG--DVERARELVEKSGLEKTRALABEYAKALDAI 416

Db 242 PLIAAQMOTGEORDLIRRSATGTSQLEQVIAIVQNSGALDYCHKRAETEERLQAL 301

QY 417 RTPESPARKALBQLTDKVLTR 438

Db 302 EILPESTYRQALVNLRLALDR 323

RESULT 8

US-09-134-001C-3250

; Sequence 3250, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 3250

; LENGTH: 357

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-3250

Query Match 15.8%; Score 346; DB 4; Length 357;

Best Local Similarity 24.8%; Pred. No. 2.4e-25;

Matches 95; Conservative 73; Mismatches 147; Indels 68; Gaps 6;

QY 65 LINPLNL-VGPENSLNLSRLLSGHPSLDTVAKYVVOSEGHKIRPLMLVLLMAQATEV 123

Db 36 VINVAKLININNEIKKVEKLEAEIISDQTLQEASFHLLSGGKRVPAFVILSGQF--- 92

QY 124 APKVQGVKEVVPVNEGLAPPEVLNDKNDMMNMRSGPLTKGTEGQTSNILASORRL 193

Db 93 -----GSNNKPSEDTYR 105

QY 184 AEITEMIHAASLLHDDVIDASSTRNAPSGNQAFGNKMAILAGDFLLGRASVALARLNP 243

Db 106 AVALLEIHMATLVHDDVIDKSRGRGLTISKWDQSTALTGNFLLAMGLKHLSEISDT 165

QY 244 EVIELLATVIANLVEGEFMOLKNTVDDAIEATATQETFDYLOKTYLKTASLIASCRAS 303

Db 166 RVHSTISKISVDVCRGELFQFQDN-----SNQITNLYLRNRKTKALLQLSTQVG 218

QY 304 ALLGGATPEVADAAAYGRNGLIAGFOIVDDMLDVTVSATDLGKPGAGADLQGLATAPALF 363

Db 219 AITSNASNDVIRKMLTGTHYVIGHSFQIIDDVLDFTSSEKKLGPVSGDLNMNGHITLPVLL 278

QY 364 AWKHAEELGPMIKRKF--DPGDERARE---LVEKSDGLEKTRALAEYAKALDAIR 417

Db 279 EMRKN-----KTFKDKISQLNPDSPQAFETCITIIROSESIBOSKQISEKILNKAINLID 334

QY 418 TTPESPARKALBQLTDKVLTRSR 440

Db 335 ELEDGPNKELFRKLIKMGSRNK 357

RESULT 9

US-09-217-609A-2

; Sequence 2, Application US/09217609A

; Patent No. 6071733

; GENERAL INFORMATION:

; APPLICANT: MURAMATSU, Masayoshi

; APPLICANT: KOIKE, Ayumi

; APPLICANT: OGURA, Kyozi

; APPLICANT: KOYAMA, Tanetoshi

; APPLICANT: SHIMIZU, Naoto

; APPLICANT: CHO, Yenwin

; TITLE OF INVENTION: prenyl Diphosphate Synthetase Genes

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Kenyon & Kenyon

; STREET: 1025 Connecticut Avenue, NW - Suite 600

; CITY: Washington

; STATE: DC

; COUNTRY: US

; ZIP: 20036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2

; SOFTWARE: WordPerfect 6.1 for Windows

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/217,609A

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/873,235

; FILING DATE: 11-Jun-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: TOFFENETTI, Judith L.

; REGISTRATION NUMBER: 39,048

; REFERENCE/DOCKET NUMBER: 10235/1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-429-1776

; TELEFAX: 202-429-0796

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 325 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-217-609A-2

Query Match 15.3%; Score 336.5; DB 3; Length 325;

Best Local Similarity 26.7%; Pred. No. 1.7e-24;

Matches 100; Conservative 59; Mismatches 148; Indels 67; Gaps 6;

QY 72 VGEMSNLSNLSRLLSGHPSLDTVAKYVVOSEGHKIRPLMLVLLMAQATEVAPKVQGW 131

Db 10 LNPYIIIEVKRELYECIQSDSEINKAAHILSSGGKRVPMFVLLSG----- 56

QY 132 KVEVPVNEGLAPPEVLNDKNDMMNMRSGPLTKGTEGQTSNILASORRLAEITEMIH 191

Db 57 -----FLNDTQKDDL-----IRTAVSLELVH 77

QY 192 AASLLHDDVIDASSTRNAPSGNQAFGNKMAILAGDFLLGRASVALARLNPVEIELLAT 251

Db 78 MASLVHDDYIDNSDMRGRNTSVHIAFDKDTAIRTGHLLARALQNTATINNSKFQIFSK 137

QY 252 VIANLVEGEFMOLKNTVDDAIEATATQETFDYLOKTYLKTASLIASCRASALLGATP 311

Db 138 TILVCFGEFDQWADRFNVPVSTA-----YLRINKTKAILIEASCHLGSLSQOLDE 190

QY 312 EVADAAYAYGRNIGLAFQIVDDMLDVTVSATDLGKPGAGADLQGLATAPALFAWKHAE- 370

DB 191 QSTYHIKQFHGCTGMSYQIIDDIDYTSDEATLTKGPVGSDIRNCHITYPYLMALIANLKEQ 250
371 ---LGMIRKRFSPDGVARELVE--KSDGLEKTRALAEYAKALDAIRTFEPSPA 424
DB 251 DDDKLEAVVGH-LTSTDDEYQYIVSQVKQYGIIEPABLLSRKYGDKAKYHLSQLODSNI 309
QY 425 KRALEQLTDKVLTR 438
DB 310 KDYLEEIEHEKMLKR 323
RESULT 10
US-08-873-235B-2
Sequence 2, Application US/08873235B
Patent No. 6174715
GENERAL INFORMATION:
APPLICANT: MURAMATSU, Masayoshi
APPLICANT: KOIKE, Ayumi
APPLICANT: OGURA, Kyozo
APPLICANT: Koyama, Tanecoshi
APPLICANT: SHIMIZU, Naoto
APPLICANT: CHO, Yewmin
TITLE OF INVENTION: Prenyl Diphosphate Synthetase Genes
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Kenyon & Kenyon
STREET: 1025 Connecticut Avenue, NW - Suite 600
City: Washington
STATE: DC
COUNTRY: US
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
SOFTWARE: Wordperfect 6.1 for windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/873,235B
FILING DATE: 11-Jun-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 154441/1996
FILING DATE: 14-Jun-1996
ATTORNEY/AGENT INFORMATION:
NAME: TOFFENETTI, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 10235/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-429-1776
TELEFAX: 202-429-0796
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-873-235B-2
Query Match 15.3%; Score 336.5; DB 4; Length 325;
Best Local Similarity 26.7%; Pred. No. 1,7e-24;
Matches 100; Conservative 59; Mismatches 148; Indels 67; Gaps 6;
DB 10 LNPFTIEBKRLYCIGDSSTIKAAHHIISGCKVRPFVLLSG----- 56
QY 132 KVEEVPVNEGAPPEVLNDKNPDMNMMNSGPLTKDGEIEGQTSNIILASORRLAEITEMIH 191
DB 57 -----FLNDTQKDL-----IRTASLELVH 77
QY 192 AASLIHDVDIDASTRRNAPSGNOAFGNKAILLAGDPLGRASVALARLNPEVIELLAT 251

DB 78 MASIVHDDYIDNSDMRGRGTSVHIAFDKDTAIRGHPILARALQNIATINNSKHFQIFSK 137
QY 252 VIANLVEGEFMQKNTVDDAIEATATQETFDYYLOKTYLKTASLIASKCRASALIGGATP 311
DB 138 TLEVOCFGEFDDMADRFPNVSFTFA-----YLRIRNRKTAIIIEASCHGALSQLODE 190
QY 312 EVADAAYAGRNGLAFOIVDDMDLTVSATDLGKRPACADQLOGLATPAPLAWMGHAE- 370
DB 191 QSTYHIKQFHGCTGMSYQIIDDIDYTSDEATLTKGPVGSDIRNCHITYPYLMALIANLKEQ 250
QY 371 ---LGMIRKRFSPDGVARELVE--KSDGLEKTRALAEYAKALDAIRTFEPSPA 424
DB 251 DDDKLEAVVGH-LTSTDDEYQYIVSQVKQYGIIEPABLLSRKYGDKAKYHLSQLODSNI 309
QY 425 KRALEQLTDKVLTR 438
DB 310 KDYLEEIEHEKMLKR 323
RESULT 11
US-09-187-050-28
Sequence 28, Application US/09187050B
Patent No. 6043072
GENERAL INFORMATION:
APPLICANT: Croteau, Rodney B
APPLICANT: Helner, Jerry
TITLE OF INVENTION: Nucleic Acids Encoding Taxus Gerard[geranyl] Diphosphate
TITLE OF INVENTION: Synthase, And Methods of Use
FILE REFERENCE: MSUR12423
CURRENT APPLICATION NUMBER: US/09/187,050B
CURRENT FILING DATE: 1998-11-05
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 28
LENGTH: 393
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Taxus GGPP
OTHER INFORMATION: synthase protein variant
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)..(393)
OTHER INFORMATION: Computer-generated Taxus GGPP synthase protein
US-09-187-050-28
Query Match 12.9%; Score 283; DB 3; Length 393;
Best Local Similarity 26.1%; Pred. No. 3.7e-19;
Matches 119; Conservative 59; Mismatches 170; Indels 108; Gaps 15;
DB 23 TLTASAPSLRLRCTPTSPSSMAAASASRLVEPPNQLINPLNLVGPMSNLTSN 82
DB 4 TAMAGTOSLOLRITVASYECNSMRSCKF-----LTPFK-----SFGVAFN 45
QY 83 IRSLLGS-----GHSPLDTVAKYVQSEGGKHIRPLMLVMAQATEVAPKVGMEKVEVP 137
DB 46 VPSLGANCEINGHLKLGSLPYKQCSVSSKSTK-----TMAQLVLAETKXEGDIBFD 100
QY 138 VNEGAPPEVLNDKNPDMNMMNSGPLTKDGEI-EGQTSNIILASORL----- 183
DB 101 FNEYMSKAVVAVALD---KATPLEYPEKIHESMRYSLLAGGRVAPALCIAACELVG 156
QY 184 -----AETEMIIAASLHDV--IDASETRRNAPSGNOAGNKAILLAGDPLGR 232
DB 157 GSQDLAMPKACMEMIHMSLIHDDLPCMDNDFRGRKFTNHKVGEDTAVIAGALLSF 216
QY 233 A---SVALARL---RNEVIELLATVIAN--LVEGEFMQKNTVDDAIEATATQETFD 282
DB 217 AFPHIIVATSKTVPSDRTLRVISELGTIGSGLVGGVVDITSGDANVD-LKTLIEWIH 275
QY 283 YIQKTYLKTASLIASKCRASALLGGATPEVADAAYAGRNGLAFOIVDDMDLTVSAT 342

Db 276 IH-----KTAVLLECSVSGGILGGATEDEIARIYRCVGLLPQVDDILDVTKSSE 329
 QY 343 DLGKPGADLQGLATAPALFAWKHAEGLPMIKRKFSDPGDVERARELVKESDGLKTR 402
 Db 330 ELGKTAGKDLTDTKATPKLM-----GLEKAK 356
 QY 403 ALAEYAKALDAIRTPESPARKALEQLTKVLR 438
 Db 357 EFAELATRAKEELSSFDQIKAAPLL-GLADYIAFR 391

RESULT 12
 US-09-187-050-27
 ; Sequence 27, Application US/09187050B
 ; Patent No. 6043072
 ; GENERAL INFORMATION:
 ; APPLICANT: Croteau, Rodney B
 ; APPLICANT: Hefner, Jerry
 ; TITLE OF INVENTION: Nucleic Acids Encoding Taxus Geranylgeranyl Diphosphate
 ; TITLE OF INVENTION: Synthase, And Methods of Use
 ; FILE REFERENCE: WSUR12423
 ; CURRENT APPLICATION NUMBER: US/09/187,050B
 ; NUMBER OF SEQ ID NOS: 34
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 27
 ; LENGTH: 393
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Taxus GGPP
 ; OTHER INFORMATION: synthase protein variant
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)..(393)
 ; OTHER INFORMATION: Computer-generated Taxus GGPP synthase protein
 ; OTHER INFORMATION: variant
 US-09-187-050-27

Query Match 12.8%; Score 280; DB 3; Length 393;
 Best Local Similarity 25.9%; Pred. No. 7.2e-19;
 Matches 118; Conservative 59; Mismatches 171; Indels 108; Gaps 15;

QY 23 TLRTASPSLRCTPTSRSSSWAAVSSASRLVEPDNPQINPLNVLGPEMSNITSN 82
 Db 4 TAAAGTQSLQLRTVASQYCNMSRCFK-----LTPFK-----SPHGVNFN 45
 QY 83 IRSILGS-----GHPSLDTAKYVQSEGKHIRPLVLLMAQATEVAPKVGHEKVEVP 137
 Db 46 VPSLAANCIIMGHLGLSLPYKQCSVSSKSTK-----TWAQLVDLAETEAEGKQIEFD 100
 QY 138 VNEGLAPPEVLNDKPNMNMNRSGPLTKDGEI-EGQTSNIIASORRL----- 183
 Db 101 FNEYMKSAAVDAALD-----KAIPLEYPEKIHESMYSLLAGGKVRPAPALCIAACELVG 156
 QY 184 -----AEITEMHAASLLHDDV--IDASETRRNAPSGNQAFGNKMAILAGDFLLGR 232
 Db 157 GSQDLAMPTACAMEMHTMSLIHDDLPCMDNDPFRGKPTNKHVFGEDTAVLAGDALLSF 216
 QY 233 A-----SVALARL-----RNPEVIELLATVIAN--LVEGEFQKNTVDDAEATQETFD 282
 Db 217 AFEHIAVATSKTVPSDRTLRISELGKTIQSGQLVGQGVDDITSEGDANDV-LKTLWEIH 275
 QY 283 YVLOKTYLKTASLIKSCRASALIGGATPEVADAAYAGNGLAFOIVDDMLDITYTSAT 342
 Db 276 IH-----KTAVLLECSVSGGILGGATEDEIARIYRCVGLLPQVDDILDVTKSSE 329
 QY 343 DLGKPGADLQGLATAPALFAWKHAEGLPMIKRKFSDPGDVERARELVKESDGLKTR 402
 Db 330 ELGKTAGKDLTDTKATPKLM-----GLEKAK 356
 QY 403 ALAEYAKALDAIRTPESPARKALEQLTKVLR 438

Db 357 EFAELATRAKEELSSFDQIKAAPLL-GLADYIAFR 391

RESULT 13
 US-09-187-050-2
 ; Sequence 2, Application US/09187050B
 ; Patent No. 6043072
 ; GENERAL INFORMATION:
 ; APPLICANT: Croteau, Rodney B
 ; APPLICANT: Hefner, Jerry
 ; TITLE OF INVENTION: Nucleic Acids Encoding Taxus Geranylgeranyl Diphosphate
 ; TITLE OF INVENTION: Synthase, And Methods of Use
 ; FILE REFERENCE: WSUR12423
 ; CURRENT APPLICATION NUMBER: US/09/187,050B
 ; NUMBER OF SEQ ID NOS: 34
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 393
 ; TYPE: PRT
 ; ORGANISM: Taxus canadensis
 US-09-187-050-2

Query Match 12.7%; Score 278; DB 3; Length 393;
 Best Local Similarity 28.9%; Pred. No. 1.1e-18;
 Matches 103; Conservative 47; Mismatches 127; Indels 80; Gaps 11;

QY 117 MAQATEVAPKVGWEKVEVPVNEGLAPPEVLNDKPNMNMNRSGPLTKDGEI-EGQTSN 175
 Db 80 MAQLVDLAETEAEGKQIEFDNFNEYMKSAAVDAALD-----KAIPLEYPEKIHESMYS 135
 QY 176 ILASORRL-----AEITEMHAASLLHDDV--IDASETRRNAP 211
 Db 136 LLAGGKVRPAPALCIAACELVGGSDQLAMPTACAMEMHTMSLIHDDLPCMDNDPFRGK 195
 QY 212 SGNAQGNKMAILAGDFLLGR-----SVALARL-----RNPEVIELLATVIAN--LVEGEF 261
 Db 196 TNHKVFGEDTAVLAGDALLSFAFEHIAVATSKTVPSDRTLRISELGKTIQSGQLVGQV 255
 QY 262 MOLKNTVDDAIEATATQETFDYVLYQTKYLTASLIKSCRASALLGGATEVADAAYAG 321
 Db 256 VDIITSEGDANDV-LKTLWEIH-----KTAVLLECSVSGGILGGATEDEIARIYR 308
 QY 322 RNGLAFQIVDDMLDITYTSATDLGKPGADLQGLATAPALFAWKHAEGLPMIKRKFSD 381
 Db 309 RCVGLLFQVDDILDVTKSSELGKTAGKDLTDTKATPKLM----- 350
 QY 382 PGDVERARELVKESDGLKTRALAEYAKALDAIRTPESPARKALEQLTKVLR 438
 Db 351 -----GLEKAFAELATRAKEELSSFDQIKAAPLL-GLADYIAFR 391

RESULT 14
 US-09-187-050-14
 ; Sequence 14, Application US/09187050B
 ; Patent No. 6043072
 ; GENERAL INFORMATION:
 ; APPLICANT: Croteau, Rodney B
 ; APPLICANT: Hefner, Jerry
 ; TITLE OF INVENTION: Nucleic Acids Encoding Taxus Geranylgeranyl Diphosphate
 ; TITLE OF INVENTION: Synthase, And Methods of Use
 ; FILE REFERENCE: WSUR12423
 ; CURRENT APPLICATION NUMBER: US/09/187,050B
 ; CURRENT FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 34
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 14
 ; LENGTH: 393
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Taxus GGPP
 ; OTHER INFORMATION: synthase protein variant

FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)..(393)
OTHER INFORMATION: Computer-generated Taxus GGPP synthase protein
OTHER INFORMATION: variant
US-09-187-050-14

Query Match 12.7%; Score 278; DB 3; Length 393;
Best Local Similarity 28.9%; Pred. No. 1,1e-18;
Matches 103; Conservative 47; Mismatches 127; Indels 80; Gaps 11;

QY 117 MAQTEVAPKQVQMEKVEVPVNEGALPPEVLNDKPNPMNMRSGPLTKDGEI-EGQTSN 175
DB 80 MAQVLDLAETEKAGKDIPEFNEFKSKAVAVDAALD---KAIPLPEPKIHESMRYS 135
QY 176 ILASQRL-----AETEMITHAASLHLDV--IDASETRRNAP 211
DB 136 LLAGKRVPRALCTIACELVGSQDLAMPITACAMEMITHSLIHDDLPCMDNDPFRGKP 195
QY 212 SGNQAFGNKRAILAGDFLLGRA---SVALARL---RNPEVIELLATVIAN--LVGEF 261
DB 196 TNHVFGEEDTAVLAGDALLSFAFEHIAVATSKVPSDRTLRLVISLKGITGSGQVLGGOV 255
QY 262 MQLKNTVDDAIEATATQETFDYVYLOKTYLKTASLIKSCASALLGGATPEVADAAYAG 321
DB 256 VDTSEGDANDV-LKTLIEWIH-----KTAVLLECSVSGILGGATDEIARIRRYA 308
QY 322 RNGLAFQIVDDMLDYVSATDLGKPAADLQGLATAPALFAWKHHAELGPMIKRKFSD 381
DB 309 RCVGLLFQVVDLIDVTKSSSELGKTAGKDLITDKATYPKLM----- 350
QY 382 PGDVERARELVEKSDGLEKTRALAEYAQKALDAIRTPESPARKALEQLTQKVLTR 438
DB 351 -----GLEKAKEFAELATRAKEELSSFDQIKAAPLL-GLADYIAFR 391

RESULT 15

US-09-187-050-16
Sequence 16, Application US/09187050B
Patent No. 6043072
GENERAL INFORMATION:
APPLICANT: Croteau, Rodney B
APPLICANT: Hefner, Jerry
TITLE OF INVENTION: Nucleic Acids Encoding Taxus Geranyl Diphosphate
FILE REFERENCE: MSUR12423
CURRENT APPLICATION NUMBER: US/09/187,050B
CURRENT FILING DATE: 1998-11-05
NUMBER OF SEQ ID NOS: 34
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 16
LENGTH: 393
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Taxus GGPP
OTHER INFORMATION: synthase protein variant
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)..(393)
OTHER INFORMATION: Computer-generated Taxus GGPP synthase protein
OTHER INFORMATION: variant
US-09-187-050-16

Query Match 12.7%; Score 278; DB 3; Length 393;
Best Local Similarity 28.9%; Pred. No. 1,1e-18;
Matches 103; Conservative 47; Mismatches 127; Indels 80; Gaps 11;

QY 117 MAQTEVAPKQVQMEKVEVPVNEGALPPEVLNDKPNPMNMRSGPLTKDGEI-EGQTSN 175
DB 80 MAQVLDLAETEKAGKDIPEFNEFKSKAVAVDAALD---KAIPLPEPKIHESMRYS 135
QY 176 ILASQRL-----AETEMITHAASLHLDV--IDASETRRNAP 211

DB 136 LLAGKRVPRALCTIACELVGSQDLAMPITACAMEMITHSLIHDDLPCMDNDPFRGKP 195
QY 212 SGNQAFGNKRAILAGDFLLGRA---SVALARL---RNPEVIELLATVIAN--LVGEF 261
DB 196 TNHVFGEEDTAVLAGDALLSFAFEHIAVATSKVPSDRTLRLVISLKGITGSGQVLGGOV 255
QY 262 MQLKNTVDDAIEATATQETFDYVYLOKTYLKTASLIKSCASALLGGATPEVADAAYAG 321
DB 256 VDTSEGDANDV-LKTLIEWIH-----KTAVLLECSVSGILGGATDEIARIRRYA 308
QY 322 RNGLAFQIVDDMLDYVSATDLGKPAADLQGLATAPALFAWKHHAELGPMIKRKFSD 381
DB 309 RCVGLLFQVVDLIDVTKSSSELGKTAGKDLITDKATYPKLM----- 350
QY 382 PGDVERARELVEKSDGLEKTRALAEYAQKALDAIRTPESPARKALEQLTQKVLTR 438
DB 351 -----GLEKAKEFAELATRAKEELSSFDQIKAAPLL-GLADYIAFR 391

Search completed: January 16, 2003, 09:50:45
Job time : 25 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:49:07 ; Search time 19 Seconds

(without alignments)
460.345 Million cell updates/sec

Title: US-09-830-111a-2

Perfect score: 2194

Sequence: 1 MASPALRBSISRSIASLR.....ESPAKALEQLDKVLTISR 440

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 120991 seqs, 19878514 residues

Total number of hits satisfying chosen parameters: 120991

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	451.5	20.6	323	10	US-09-815-242-10321
2	451.5	20.6	323	10	US-09-815-242-13734
3	445.5	20.3	329	10	US-09-815-242-11142
4	405	18.5	322	10	US-09-815-242-12031
5	371.5	16.9	332	10	US-09-815-242-12603
6	367.5	16.8	312	10	US-09-815-242-5315
7	347	15.8	326	10	US-09-815-242-10573
8	340	15.5	350	9	US-09-738-626-4030
9	339	15.5	273	10	US-09-925-637-66
10	302.5	13.8	307	10	US-09-815-242-11306
11	272	12.4	367	9	US-10-108-915-26
12	260	11.9	299	10	US-09-815-242-10069
13	255.5	11.6	350	9	US-10-108-915-22
14	254.5	11.6	232	9	US-10-108-915-14
15	253	11.5	369	9	US-10-108-915-18
16	252	11.5	377	10	US-09-934-778-2
17	251.5	11.5	293	10	US-09-815-242-10630
18	247.5	11.3	316	9	US-10-108-915-45
19	247	11.3	371	9	US-09-738-626-5880

20	245.5	11.2	291	10	US-09-815-242-13273	Sequence 13273, A
21	241	11.0	295	10	US-09-815-242-11971	Sequence 11971, A
22	240	10.9	291	10	US-09-815-242-13597	Sequence 13597, A
23	237.5	10.8	299	10	US-09-815-242-14084	Sequence 14084, A
24	235.5	10.7	228	9	US-10-108-915-12	Sequence 12, Appl
25	226.5	10.3	345	9	US-10-108-915-44	Sequence 44, Appl
26	212	9.7	342	9	US-10-108-915-34	Sequence 34, Appl
27	202.5	9.2	209	9	US-10-108-915-30	Sequence 30, Appl
28	193	8.8	303	9	US-09-941-947A-26	Sequence 26, Appl
29	191.5	8.7	295	10	US-09-815-242-11239	Sequence 11239, A
30	189	8.6	287	10	US-09-925-637-64	Sequence 64, Appl
31	189	8.6	293	10	US-09-815-242-12583	Sequence 12583, A
32	187.5	8.5	297	9	US-09-941-947A-20	Sequence 20, Appl
33	187.5	8.5	297	10	US-09-934-903-14	Sequence 14, Appl
34	187.5	8.5	297	10	US-09-934-868-72	Sequence 72, Appl
35	185.5	8.5	382	9	US-09-738-626-4197	Sequence 4197, Ap
36	184.5	8.4	288	10	US-09-815-242-5239	Sequence 5239, Ap
37	182	8.3	326	9	US-10-108-915-46	Sequence 46, Appl
38	180.5	8.2	342	9	US-10-108-915-38	Sequence 38, Appl
39	179.5	8.2	133	9	US-10-108-915-4	Sequence 4, Appl
40	170	7.7	303	10	US-09-815-242-11382	Sequence 11382, A
41	164.5	7.5	128	9	US-10-108-915-28	Sequence 28, Appl
42	158	7.2	295	10	US-09-547-267-1	Sequence 1, Appl
43	156	7.1	303	10	US-09-815-242-11547	Sequence 11547, A
44	142.5	6.5	304	10	US-09-925-302-614	Sequence 614, App
45	141	6.4	316	9	US-10-108-915-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-815-242-10321
; Sequence 10321, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Hasebeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Cart, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10321
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10321

Query Match 20.6%; Score 451.5; DB 10; Length 323;
Best Local Similarity 31.0%; Pred. No. 1.2e-30;

Matches	116;	Conservative	52;	Mismatches	143;	Indels	63;	Gaps	4;
QY	71	LVGPEMSNLTSNIRSLISLGGHPSLDTVAKYVYQSEGGKHIRPLMLVILMAQATEVAPKVGCM	130						
Db	9	LTAQGMAGNAAILLEQLNSDVOLINLGGYIYISGGSKRIRPMIAVLAAAV-----GY	61						
QY	131	EKVEVPEVNEGGLAPPEVLNDKNPMDMMNRSGPLTKDGEIEGOTSNIILASORRLAETTEMI	190						
Db	62	EGNAHYTI-----AALIFRI	76						
QY	191	HAASLLHDDVIDASETRNRAPSGNOAFGKKMAILLADPELLGASVALAARLNRPEVIELLA	250						
Db	77	HTATLLHDVDVDESDMRKATANAAFGNAASVLDGFYTPRAFGWMTSLGSLKYLEVMS	136						
QY	251	TVIANLVSEEPQGLKTVDDAIEATRAQTGTFPDIYLOKTLKTLASLLAKSCRSALLGGTF	310						
Db	137	EAVNVIAGEBEVQLTMVNVNDPDI-----TEENYMKVIYSKTARLFEEAAQCSGIIAGCT	189						
QY	311	PEVADAAPAYKGNLGLAFQIVDDMLDYTVSAYSDILGKRPADILQLGLATAPALPAMKCHNE	370						
Db	190	PEEEKGLDQYGYLTLGTAFOILLDDLDYNADGBQLGNVGDDENEGFPPLPLHAMHHGTP	249						
QY	371	LGPMLKKKSSDGD-----VERARELVKSDGLEKTRALAEVQAOKLDAIRTFPESPAPK	426						
Db	250	EOAQMIKRAIEQNGRHLLPEVLEAMNACGSLWTRQRAEEEDAKIAALQVLPDTPWRE	309						
QY	427	ALLEQLTDKVLTRSR	440						
Db	310	ALITGLAHIAVQDR	323						

```

RESULT 2
US-09-815-242-13734
Sequence 13734, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zykink, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13734
LENGTH: 323
TYPE: PRT
ORGANISM: Salmonella typhi
US-09-815-242-13734

```

Matches	115;	Conservative	54;	Mismatches	142;	Indels	63;	Gaps	5;
QY	71	LVGPEMSNLTSNIRSLGSGHPSLDVTVAKYVYQSEKHIRPLMLVLLMAQATEVARKVQGM	130						
Db	9	LTAGQAGVGNATITELQNLMSDVQLINQLGYIILISGGKRIIRPMIAVLAAV-----GY	61						
QY	131	EKVVEPVNNEGGLAPEVILNDKNPDMNMNRSGPLTYDGEIEGOTSNIIASORRLAEITEMI	190						
Db	62	QGNAAHTI-----ALIEFI	76						
QY	191	HAASLIHDDVIDASTERNRNPASGNOAFGNKAAIILAGDPLGLASVALARLRNPVETIELLA	250						
Db	77	HTATLIHDDVDVDESDMRGKATANAAFGNAASVLTGDEITYRAFGQMTSLGSLKYLEVMS	136						
QY	251	TVIANLVGGEFMQLKNYDDAEATATGETPDIYIQLKTYLTASLIASCRASALLGAT	310						
Db	137	EAVNVIAEGEVLQILNNVNDPDI-----TEENVRYVYSKTARVLFENAAQCSGLIAGCT	189						
QY	311	PEVAAAAVAYGSGNIGLAFQIYDDMDIYVSATDILGKPGAGDLQGLATAPALFAMKHHA-	369						
Db	190	PEQEGKLODYGVYIGTAFQILIDLLDYADGEHILGKNVGDLDNECKPTPLIHHMRHGT	249						
QY	370	ELGPMIKKKFSPGD---VERARELVEKSDGLEKTRALAEVAAQKALDAIRTFPSPARK	426						
Db	250	EESAMIRFAIEGNGRHLLEPYLEAMTTCGSLSEWTRORAEEDAKAISALQILPPTPARE	309						
QY	427	ALAEQLTDKVLTRSR	440						
Db	310	ALIGLAHIAVQDR	323						

```

RESULT 3
US-09-815-242-11142
; Sequence 11142, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: EILTRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/266,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11142
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-11142

```

Query Match	20.6%;	Score 451.5;	DB 10;	Length 323;
Best Local Similarity	30.7%;	Pred. No. 1.2e-30;		

Query Match	20.3%;	Score 445.5;	DB 10;	Length 329;
Best Local Similarity	32.5%;	Pred. No. 3.8e-30;		

Matches 108; Conservative 56; Mismatches 146; Indels 78; Gaps 9;

QY 59 PDPNQLINPLNVGPEMSNLTNIRSLGSHPSLDTVAKYVYQSGKHIRPLMLVILMA 118
 Db 9 PDVERSLTETCELIEQTLHTRNEDIEE-----ALVTLAR-----SGGKLLRPAPFFLFA 57

QY 119 QATEVAPKVGWEKVEVPVNEGLAPLEVLDNDKPNPMNMRSGPLTKDGEIGOTSNILA 178
 Db 58 QL-----GGEKEKQOOLL 71

QY 179 SORRLABITEMIHAASLLHDDVIDASETRRNAPSGNOAFGNKMAILAGDFLLGRASVALA 238
 Db 72 --KIAASLEILHWATLIHDDIIDDSPLRGAVTIQSGYKQVAVYTGDLTLFTFFFLIA 128

QY 239 RLRY-PEVIELLATVIANLVEGEFMQKNTVDDDAIEATATQETPDYLYQKTYLKTASLIA 297
 Db 129 DAMNGSEFMKINAOQMKRLLLGELDOMSHRFDRRMSIPA-----YLRVNGKTAEFLS 181

QY 298 KSCRASALLGGATPEVADAAYVAGRNIGLAFQIVDDMLDVTVSATDLGKPGADLQGLA 357
 Db 182 LSCLEGAYFGHSSKEVQRLAKRIGRHIGIAPQVYDDILDYADTETLKKPALEDLSQGVY 241

QY 358 TAPALFAWKHAEIGPMIKRKFSPG---DVERARE---LVEKSDGLEKTRALAEVYAKQ 411
 Db 242 TLPLLFAY---QAAPDVFSYLDKGRAITLLEAAEVAALVNDYHGVTAEQAPAKKYTNK 297

QY 412 ALDAIRTFPSPARKALEQLTKDVLTRS 439
 Db 298 AITDIOQLPDGTAKETLLSLTELLHRS 325

RESULT 8

US-09-738-626-4030
 ; Sequence 4030, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIALI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OKAZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 4030
 ; LENGTH: 350
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-4030

Query Match 15.5%; Score 340; DB 9; Length 350;
 Best Local Similarity 27.4%; Pred. No. 3.4e-21;
 Matches 115; Conservative 58; Mismatches 155; Indels 92; Gaps 14;

QY 35 RCTPT-SRPSSSMAAVSSASRLVEPPD-----NQPLINPLNVGPEMSNLTNIRSL 87
 Db 5 RTVPTRSHGLGKEGVSTTGASQVEFGDPDLTARINDAMVQVEELLHTLS----- 54

QY 88 GSGHPSLDTVAKYVYQSGKHIRPLMLVLLMAQAATEVAPKVGWEKVEVPVNEGLAPPEV 147

Db 55 -SGEDFLVDIVMLHTRAGGRFRPFALL---ASEFGEK-----PSE----- 93

QY 148 LNDKNPDMNMRSGPLTKDGEIGOTSNILASQRRRLAEITEMIHAASLLHDDVIDASETR 207
 Db 94 -----NVI-----KAAVVEITHLATLYHDDVMDSEASMR 122

QY 208 RNAPSGNOAFGNKMAILAGDFLLGRASVALARLNPEVIELLATVIANLVEGEFMQKNT 267
 Db 123 RGVPFSANARWNSVAILAGDILLAHASGLMSQL-GTDTVAHFAETFGELVTG---QMRET 178

QY 268 VDDAIEATATQETPDYLYQKTYLKTASLIAKSCRASALLGGATPEVADAAYVAGRNIGLA 327
 Db 179 V-----GPRDTPDIEHTNVIREKTGVLIASAGYLGMAHGAAPAHIDALKNFGAAGMI 233

QY 328 FOIVDDMLDVTVSATDLGKPGADLQGLATAPALFAWKH-----AELGPMIKRKFSPDG 383
 Db 234 FOIVDDIIDIFETHESGKTPGTDLREGVFTLPVLVALREDTPVGAEFLRDLITGPLEDDE 293

QY 384 DVERARELVEKSDGLEKTRALAEY-----AQKALDAIRTFPSPARKALEQLTKDVLTR 438
 Db 294 TVNHVLELLSQSGG--RQAALDEVRYMDIANAELORL---PDSTVKEALRNLATFTVKR 348

RESULT 9
 US-09-925-637-66
 ; Sequence 66, Application US/09925637
 ; Patent No. US2002010338A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Choi
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
 ; FILE REFERENCE: PB560
 ; CURRENT APPLICATION NUMBER: US/09/925,637
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/23773
 ; PRIOR FILING DATE: 2000-08-31
 ; PRIOR APPLICATION NUMBER: US 60/151,933
 ; PRIOR FILING DATE: 1999-09-01
 ; PRIOR APPLICATION NUMBER: US 08/781,986
 ; PRIOR FILING DATE: 1997-01-03
 ; PRIOR APPLICATION NUMBER: US 08/956,171
 ; PRIOR FILING DATE: 1997-10-20
 ; PRIOR APPLICATION NUMBER: US 60/009,861
 ; PRIOR FILING DATE: 1996-01-06
 ; NUMBER OF SEQ ID NOS: 74
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 66
 ; LENGTH: 273
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (361)..(261)
 ; OTHER INFORMATION: xaa equals any amino acid
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (271)..(271)
 ; OTHER INFORMATION: xaa equals any amino acid
 US-09-925-637-66

Query Match 15.5%; Score 339; DB 10; Length 273;
 Best Local Similarity 30.6%; Pred. No. 2.9e-21;
 Matches 85; Conservative 56; Mismatches 121; Indels 16; Gaps 3;

QY 160 SGPLTKDGEIGOTSNILASQRRRLAEITEMIHAASLLHDDVIDASETRRNAPSGNOAFGN 219
 Db 5 SSQFGKDEQTSQTYQV-----AVALEIHWATLHVHDDVIDKSKRRKGLTISKWDQ 57

QY 220 KMAILAGDFLLGRASVALARLNPEVIELLATVIANLVEGEFMQKNTVDDAIEATATQE 279
 Db 58 TTAILGNFLLALGLEHMAVKDNVRHQLSSESIVDVCRGELFQFDQFN-----SQ 110

QY 280 TFDYLYQKTYLKTASLIAKSCRASALLGGATPEVADAAYAGRNIGLAFOIVDDMLDVT 339


```

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10069
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-815-242-10069

```

```

Query Match 11.9%; Score 260; DB 10; Length 299;
Best Local Similarity 27.5%; Pred. No. 1.6e-14;
Matches 114; Conservative 45; Mismatches 104; Indels 152; Gaps 19;

QY 49 AVSSASRLVDPDP--NOPLINPLNLVGPMSNLTSNIRSLGSGHPSLDTVAKYVQSEG 106
DB 14 ANQALSRIAPLPQNTPVVE-----TMQYGALLG-----G 44

QY 107 KHIRPLMVLMAQATEVAPKVGQWKEVVEPVNEGLAPPEVLNDKNPMMNMRSGPLTKD 166
DB 45 KRLAPFLV---YAT----- 55

QY 167 GEISGQSNILASQRRRLAEITEMIHAASLLHDDV--IDASETRNAPSNGQAFGNKMAIL 224
DB 56 GHMFGVSTNTLDAP---AAAVECIHAYSLIHDDLPAMDDDLRLRGLPTCHVKFGEANAIL 112

QY 225 AGDFLLGRASVALARLRNPEV-----IELL-----ATVIANLVEGEFQML-----KNTVDD 270
DB 113 AGDALQTLASILSDADMPVEDSDRISMISELASAGIACMGCGQALDLDAEGKHVPLD 172

QY 271 AIEATATQETFDYLYQTYLKTASLIKSCRASALLG-----ATPEVADAAYAGRNGLG 325
DB 173 ALERIHRH-----KTGALIRAAVRLGALSAGDKGRALP-VLD---KYAESIG 216

QY 326 LAFOIVDDMLDYTVSATDLGKPGADILQGLATAPALPAWKHAEELGPMIKRKFSDPGDV 385
DB 217 LAFOVQDDILDVVGDTATLGKROGADQOLGKSTYPAL-----LLEQARK----- 261

QY 386 ERARELVKSDGLEKTRALAEVYAKALDAIRTPESPARKALEQDITDKVLTISR 440
DB 262 -KARDLID--DARQLKQLAEQ-----SLDT-----SALEALADYIIQRNK 299

RESULT 13
US-10-108-915-22
; Sequence 22, Application US/10108915
; Patent No. US20020177204A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Shen, Jennie
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: Geranylgeranyl Pyrophosphate Synthases
; FILE REFERENCE: BB1286 US NA
; CURRENT APPLICATION NUMBER: US/10/108,915
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/452,238
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,592
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-108-915-22

Query Match 11.6%; Score 255.5; DB 9; Length 350;
Best Local Similarity 23.4%; Pred. No. 4.7e-14;
Matches 94; Conservative 59; Mismatches 141; Indels 107; Gaps 11;

QY 37 TPTSREPSSWAAVSSASRLVDPDPNOPLINPLNLVGPMSNLTSNIRSLGSGHP-SLD 95
DB 33 TVTKRRAFSLSAVLT-----VETEKKPPIFDKKNYMLSKASAVNKGLDDSVSLREPKKIH 87

QY 96 TVAKYVQSEGKHIRPLMVLMAQATEVAPKVGQWKEVVEPVNEGLAPPEVLNDKNPDM 155
DB 88 EAMRYSLLAGGKVRPVLV-----ACELVGGHEATAMPA----- 123

QY 156 MNMRSGPLTKDGEIGQTSNILASQRRRLAEITEMIHAASLLHDDV--IDASETRNAPS 213
DB 124 -----ACALEMIHTMSLIHDDLPCMDNDLRRGKPTN 155

QY 214 NQAFGNKMAILAGDFLLGRASVAL--ARLRNPEVIELLATV--IANLVEGEFQMLKNTVDD 270
DB 156 HTVFGEDVAVLAGDALLAFEHIAASTRGASAPRILRAIGELARSIGSEGLVAGQVVDI 215

QY 271 AIEATATQETFDYLYQTYL-----KTASLIKSCRASALLGSGATPEVADAAYAGRNGLG 326
DB 216 NSBGLA-----DVGLERLEFHHVHKTAALLEGAVVLGAILGGTDDDEVEKLRKFARYIGL 270

QY 327 AFOIVDDMLDYTVSATDLGKPGADILQGLATAPALPAWKHAEELGPMIKRKFSDPGDVE 386
DB 271 LFQVQDDILDVTKSSQELGKTAGKDLVADKVTYPKLL----- 307

QY 387 RARELVKSDGLEKTRALAEVYAKALDAIRTPESPARKA 427
DB 308 -----GIBKSFEFAKLNKDAQDQLAGF--DFVKAA 336

RESULT 14
US-10-108-915-14
; Sequence 14, Application US/10108915
; Patent No. US20020177204A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Shen, Jennie
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: Geranylgeranyl Pyrophosphate Synthases
; FILE REFERENCE: BB1286 US NA
; CURRENT APPLICATION NUMBER: US/10/108,915
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/452,238
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,592
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 232
; TYPE: PRT

```

ORGANISM: Oryza sativa
US-10-108-915-14

Query Match 11.6%; Score 254.5; DB 9; Length 232;
Best Local Similarity 29.3%; Pred. No. 3.1e-14;
Matches 75; Conservative 39; Mismatches 93; Indels 49; Gaps 6;

QY 187 TEMHAASLHDDV--IDASETRNAPSNGOAFGNKAILAGDFLGRASVALARL-----240
DB 1 TSIQHTMSLVHDDLPCHDDDDLRGKPEPTCHVVGEPFAVLTGALLSLSPHMAARFDSYP 60

QY 241 -----RNPEVIELLAVIANLVGEFMOJKNVTVDALIEATQETPDY-YLOKTYL-KT 232
DB 61 PDIDADKPARVRAISELARCISEGIVAGQVVD--LEMTGSTETVPLERLEIYHLHT 118

QY 293 ASLIAKSCRASALIGATPEVADAAYVGRNLGLAFOIVDMLDYTVSATDLGKPAQADL 352
DB 119 ALLENAVIVGAILGGSDGQIESLRMYARSIGLLFQVVDIDLVTKSEELGKTAKDL 178

QY 353 QLGATAPALFAMKHAELGPMIKRKSPPGDVERARELYEKSDGLEKTRALAEVQA 412
DB 179 ASDKTYPKLL-----GLEKSRFEAKLLSDA 205

QY 413 LDAIRTFESPARKAL 428
DB 206 REQLSGFDQETAPLL 221

RESULT 15
US-10-108-915-18

Sequence 18, Application US/10108915
Patent No. US20020177204A1
GENERAL INFORMATION:

APPLICANT: Cahoon, Rebecca E.
APPLICANT: Shen, Jennie
APPLICANT: Williams, Mark
TITLE OF INVENTION: Geranylgeranyl Pyrophosphate Synthases
FILE REFERENCE: B01286 US NA
CURRENT APPLICATION NUMBER: US/10/108,915
CURRENT FILING DATE: 2002-03-28
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/452,238
PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,592
PRIOR FILING DATE: EARLIER FILING DATE: 1998-December-02
NUMBER OF SEQ ID NOS: 46
SOFTWARE: Microsoft Office 97
SEQ ID NO 18
LENGTH: 369
TYPE: PRT
ORGANISM: Glycine max
US-10-108-915-18

Query Match 11.5%; Score 253; DB 9; Length 369;
Best Local Similarity 25.2%; Pred. No. 8.3e-14;
Matches 105; Conservative 54; Mismatches 163; Indels 94; Gaps 11;

QY 59 PDPNQLINPLNIVGPEMSNLTSNIRSLGSHPSL-----DTVAKYVQS 104
DB 10 PRPTSMNTSHTLPPHTLTTLTKLASGTPKLISFPLVASVPTKEHTVTTQEIQ- 68

QY 105 EGKHIRPLMLLMAQATEVAKVQGEKVEVVPNEGILAPPEVLNDKNDPMNMRSGLT 164
DB 69 -----LQDPPLNDFPKGYMIAKAHTVNAQALDAIALRDPKHIOAMRYSLLA 115

QY 165 KDG-----EIGQTSNIIASQRLAETEMTHAASLHDDV--IDASETRNAP 211
DB 116 GGRVRPVLCTIACELVGTIE--ATAIPACAVEMHTMSLIHDDLPCHNDNDLRGKP 172

QY 212 SGNQAFGNKAILAGDFLGRASVALARL-----SVALARLNPPEVIELLATVIA--NLVEGEFM 262
DB 173 TNHKVYGEDVAVLAGDALLAFAPFHVAASTEGVSPSRVVARAIGELAKSIGTBGLVAGQV 232

QY 263 QLNKTVDDAIEATATQETPDY-YLOKTYLKTASLIAKSCRASALIGATPEVADAAYAGR 322

DB 233 D-----IDSEGVANNGLETFLEFI---HYKTAALLAAVVLGAIVGGSDDEVEKLRKFAR 285
QY 323 NGLAFQIVDMLDYTVSATDLGKPAQADLQGLATAPALFAMKHAELGPMIKRKSFP 382
DB 286 CIGLFQVVDIDLVTKSEELGKTAKDLVADKTYPKLL----- 326

QY 383 GDEVERARELYEKSDGLEKTRALAEVQAQALDAIRTFESPARKALEQITDKVLT 438
DB 327 -----GIDKSKEFAQELLDAKEQLSGF-DPPKAAPLFAITVYIAYR 367

Search completed: January 16, 2003, 09:56:38
Job time : 21 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1023	46.6	473	1	XUBVTP	trans-pentaprenylt
2	910.5	41.5	378	2	JC5429	di-trans,poly-cis-
3	823.5	37.5	356	2	T43193	trans-pentaprenylt
4	676.5	30.8	323	2	T56966	geranylgeranyl pyr
5	664.5	30.3	323	2	A40433	prephytolene pyroph
6	664.5	30.3	323	2	T06874	probable prenyl tr
7	661.5	30.2	323	2	AH1818	solaneyl diphosph
8	604.5	27.6	323	2	S73189	prenyl transferase
9	602.5	27.5	377	2	G87775	protein C24Al1.9 {
10	583.5	26.6	330	2	F96813	hypothetical prote
11	574.5	26.2	379	2	C83063	prenyl transferase
12	561	25.6	297	2	A84759	probable trans-pre
13	451.5	20.6	323	2	E65109	octaprenyl-diphosp
14	451.5	20.6	323	2	B91137	octaprenyl diphosp
15	451.5	20.6	323	2	E85982	octaprenyl diphosp
16	445.5	20.6	323	2	AH0903	octaprenyl-diphosp
17	445.5	20.3	329	2	I64160	hypothetical prote
18	424.5	19.3	348	2	A82322	octaprenyl-diphosp
19	420.5	19.2	362	2	AF3443	geranyltranstransf
20	418.5	19.1	323	2	AB0427	octaprenyl-diphosp
21	418.5	19.1	338	2	AD2654	octaprenyl-diphosp
22	418.5	19.1	358	2	AD7435	decaprenyl diphosp
23	415	18.9	338	2	H87510	polyprenyl synthet
24	411.5	18.8	320	2	B75457	polyprenyl synthas
25	405	18.5	348	2	E69630	heptaprenyl diphos
26	405	18.5	322	2	C83075	octaprenyl-diphosp
27	402.5	18.3	335	2	E70549	probable heptapren
28	395	18.0	332	2	AD2688	octaprenyl-diphosp
29	378	17.2	324	2	A81213	octaprenyl-diphosp

```

Db      4  RSGAAHIKILSRRCFRKSFPAVALNMAASKLVT-----KILMNPDISVSKEMTTLAKN  59
Oy      83  ISSLGSGHPSLDYAKVYVQSEGHIRPLMVLMLQAQTEVAPKQ-----GHKRYEV  136
Db      60  IVALIGSHPVLNKVTSYFEEFGKVPRLVLLLSRALSEIPTMERHMLKIDKSDVED  119
Oy     137  PV-----NEGAPPEVLNDKVPDMNNRSGLPTDGEIEGOTSNIILASQRL  183
Db     120  PLYKSPQNQLQRPASSISPLHILGHKRP-LNPLTKGEPRLPEETFPKORILPKQRLL  178
Oy     184  AETTEMHAASLLHDVIDASETRRNAPSGNOAFGNKWAIIAGDPLLGASVALARLNP  243
Db     179  AEIVEMIHFTASLLHDVIDHSDTRRGPRPSGNAAFTNKAVIAGDEFLGATVISRLHP  238
Oy     244  EYIELLATYIANLVBESEFQOLKTYDDAEAT-----  275
Db     239  EYVELMNSIANLVBESEFQOLKNTSIDADIDITENGHKLPLVPSKLEVEKEHDFVPSRQ  298
Oy     276  -----ATOETFPDYQKTYTKLTASLIASCRASALLGGATEPEVADAAYAGRNLG  326
Db     299  OGLQLSHDOIETAPERYIHKTYLKTALISKSCRAALISASPAVIDDECDFGRNLGI  358
Oy     327  AFOIVDDMLDYVSATDLGKPGAGADIQGLATAPALPAWKHHAAELGPMIKRRKSDPGVE  386
Db     359  CFQIYDDMDLDFYVSGDKLGKPGAGADIKLGIAIAPVLPFAKEDPSPGLISRNFSRGEVE  418
Oy     387  RARELVKSDGLEKTRALAEVYAOKALDAIR--TFESPARKLLEQDTDKVLTLSRR  440
Db     419  KIDSIRRLHNGIAKTIILAEYRDKALQWLROSLPEPSDRASLLEPLTNSILTRRK  473

```

[illegible]

```

Db      92 ID-----RSVSGEKYIDDDDLKS-----FSGQLPESQKLAAQI 125
QY      187 TEMTHAASLLHDDVIIDASETRRRNAPSGNOAFGNKMAIIACDFLLGRASVALARLRPEVT 246
      126 TEMTHIASLLHDDVIDIHANVRKSGPSSNVAFGNRRSIIAGNFILIRASTMAALRLRPQVT 185
QY      247 ELMTVTIANIVESGFPMOLKTNVIDAIEATATQETPFDDYLLQKTYLKTASLAKSCRASALL 306
Db      186 ELMTAVIADLVKRGFFLLQKNTMDS-SLEIKOSNFDYIIKESFSLKASTLSKCKASTIL 244
QY      307 GGAETPEVADAAYAAAGRNIGLAFQIIVDDMLDYTSATDLGKPGAGADQLQGLATAPALPAMK 366
Db      245 GQCSPTVATANGVEYGRGICGTAFQIUMDDVDLYTISKDDTLTGKAAGDILKLGATAPVLPAMK 304
QY      367 HHAEIIGPMIKRPSDPCQDVERARELVKESDGLKTRALAEVYAKALDAIRTPESPAPK 426
Db      305 KYPEIGAMIVNRFMHPEDIQRASLIVECTDAIEQITWIAKEYIKKAKDSLCLCPDSPARK 364
QY      427 ALLEQLTDKVLRRSR 440
Db      365 ALPALADRVITRRK 378

```

```

RESULT 3
T43193
trans-pentaprenyltransferrase homolog - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 11-Jan-2000 #sequence_version 11-Jan-2000 #text_change 21-Jul-2000
C:Accession: T43193
R:Yoshida, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA:Ref. 4, 369-369, 1997
A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A:Reference number: Z17323; MUID:98162722; PMID:9501991
A:Accession: T43193
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-356 <YOS>
A:Cross-references: EMBL:D89265; NID:G1749737; PIDN:BAAL1392.1; PID:G1749738
A:Experimental source: Strain PR745
C:Superfamily: prenyl transferase A

Query Match          37.5%; Score 823.5; DB 2; Length 356;
Best Local Similarity 50.0%; Pred. No. 5.5e-50;
Matches 175; Conservative 56; Mismatches 92; Indels 27; Gaps 4;

Oy 67 NPLNVGEMENLTGNSILSGCHPSLDYAKYVJGSEGHIRPLWVLTMAQATEVAPK 126
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 32 NASHIKMELQISGICIGMLNSSEFLECSKYTTAOGQMRSPVLVLSKATSL--- 88

Oy 127 VQGMKEVVEPVNBSGLAPEVLNDKNPDMNMRSGPLTKDGEIGQTSNIIASQRRLAET 186
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 FPGIDR-----SVYGDKYTDDDLRS-----FSTQIIPSQLRLAQI 125

Oy 187 TEMTHAASLLHDDVDVDAETRRNAPSGNQAQGNKALLIAGPFLGRASVALARLNPEVI 246
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 126 TEMTHIASLLHDDVDVDAHNVRRGSSSNVAGNRSSIIAGNFIARASTVWARLENPVT 185

Oy 247 ELLATVAVNLVEGSEMOQKNTVDDAIEATATQEFDFYLOQTYLKTASTLAKSCRASALL 306
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 186 ELLATVAVNDLVRGELFQJAKNTMDPS-SLEIRQSNFDYIIERSPLKTASLISKSCAASITLL 244

Oy 307 GGATPEVADAAVAYAGRNIGLAFQIVDDMLDYTSATDLGKPDAGDLOGLATAPALPFAMK 366
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 245 GGCSPTVATAAGEGRCIGTAFQJLMDVDLVYTSKDDTLGKAAGADVLKGLATAPALPFAMK 304

Oy 367 HHAELGPIIKKKFSDGPVEPARBELVEKSDGLEKTRALAEYAQALDAI 416
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 305 KYPELGAMIVNRFNHPISDIOARSRIVECCSDAIEOPTWAIYIKKRPSPV 354

```

geranylgeranyl pyrophosphate synthase crtE homolog - *Synechocystis* sp. (strain PCC 6803)
N;Alternate names: Hypothetical protein slr0611
C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Accession: S76966; S74427
C:Date: 14-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 21-Jul-2000
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shampo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76966
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-67 <KAN>
A:Cross-references: EMBL:D90917; GB:AB001339; NID:g1653836; PIDN:BAA20922.1; PID:g342616
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
A:Accession: S74427
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 68-323 <KAW>
A:Cross-references: EMBL:D90899; GB:AB001339; NID:g1651650; PIDN:BAA16579.1; PID:g16516516
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: crtE
C:Superfamily: prenyl transferase A

Query Match 30.8%; Score 676.5; DB 2; Length 323;
Best Local Similarity 40.8%; Pred. No. 8.7e-40;
Matches 153; Conservative 63; Mismatches 102; Indels 59; Gaps 4;

QY 65 LINPLNLVGP---EMSNLTNIRSLGSHPSLDTVAKYVQSEGHIRPLMWLLMAQAT 121
Db 1 MISTTSLFAPVDQDLRLTDLNKLVGARHPILGAAAEHLFEAGGKVRPAIVLLVSRAT 60
QY 122 EVAPKVGWEKVEVPVNEGGLAPPEVLNDKNPMMNMRSGELTKDGEIGQTSNIIASQR 181
Db 61 -----LLDQ-----ELTARHR 71
QY 182 RLAEITTEMIHAASLLHDDVIDASSTRNAPSGNOAFGNKMAILAGDFLGRASVALARLR 241
Db 72 RLAEITTEMIHTASLVHDDVDVEADLRNRPVTNSLFDNRVAVLAGDFLFAQSSWYLANLD 131
QY 242 NPEVIELLATVIANLVEGEFMQLKNTVDDATEATQETFDYLOKTYLKTASLIAKSCR 301
Db 132 NLEVVKLLSEVIRDFAGEILQSNRFD-----TDTLETYLEKSYFTASLIANSK 184
QY 302 ASALLGGATPEVDAAYAGRNGLAFQIVDDMLDVTYSATDLGKPGAGDLQGLATAPA 361
Db 185 AAGVLDAPRDVCDHLVEYKGHLGAFQIVDDILDFTSPTFVLGKPGAGSDLLISGNIITAPA 244
QY 362 LFAWKHHAELCPMTKRKFSDFCDVERARELVKSDGLEKTRALAEYAAQKALDAIRTFPE 421
Db 245 LFAWEKYPLLGKLIEREFAQGDLEQALELVEQDGGIRRSRELAANOQAQLARQHLSVLEM 304
QY 422 SPARKALEQLTDKVLTR 438
Db 305 SAPRESLLELVVDYVLR 321

RESULT 5
A40433
prephytone pyrophosphatase dehydrogenase (crtE) homolog - *Cyanophora paradoxa*
C:Species: *Cyanophora paradoxa*
C:Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 24-Sep-1999
C:Accession: A40433
R:Michalowski, C. B.; Loeffelhardt, W.; Bohnert, H. J.
J. Biol. Chem. 266, 11866-11870, 1991
A:Title: An ORF323 with homology to crtE, specifying prephytone pyrophosphate dehydroge
A:Reference number: A40433; MUID:91268060; PMID:1711042
A:Accession: A40433
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-323 <MIC>
A/Cross-references: GB:M61174; GB:M61174; NID:g336638; PIDN:AAA65472.1; PID:g336639
C:Superfamily: prenyl transferase A

Query Match 30.3%; Score 664.5; DB 2; Length 323;
Best Local Similarity 39.5%; Pred. No. 6e-39;
Matches 147; Conservative 69; Mismatches 97; Indels 59; Gaps 4;

Qy 70 NLVGP---EMSNLTNSIRSLGSGHPISLDTVAKYVVQSSEGHKIRPLMVLMLMAOATEVAPK 126
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 6 NILAPVENELDLTKNLKLVGSCHPILSAASEHLFSASGRPRPAIVLLLSKA-----59
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 127 VQGKEKVVEVPNEGGLAPEVLNDKNPDMMNRSGPLTKDGEIEGTSTNILASORRLAEI 186
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 60 -----TWNEI-----ITSRRHRLAEI 76
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 187 TEMIHAAAILHHDDVIDASETTRNAPSNGAFGNKMMAILAGDFLLGRASVALARLNPEVI 246
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 77 TEIIHTASLVHDLDIDESDVRRGIPTVDHFDGFKIAILLADGFLLFAOSSWYLANLESLEUV 136
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 247 ELLATVINANLVEGESFQLNKNTVDDDATEATQTFTFYLYQKTLYLKTSLIAKSQRASALL 306
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 137 KLIISKVIITDFAGEITR-----GLNQKFVTLTLEEYLEKSYKTASILAAASKAAAAAL 189
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 307 GGATPEVADAAYAGRNLGLAQFIQVDDMDLDYTVSATDLGPKAGADLOLGUATPALPAWK 366
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 190 SHVDULTVANLNYGRHLGLAQFIQVDDILDFTSSTEELGKPSCSDLKKGNLTAPVLFAL 249
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 367 HHAEGLGMPIKRFPDPGDVERABELVKSDGLEKTRALAEEYAOKALDAIRTTPESPARK 426
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 250 QNSELIPTIQRFSPDKPFYTLQIVEETKAIEKRELAMEHAQAIVAICLENLPSSSKES 309
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 427 ALSQLTDKVLTR 438
 || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 310 AKKLTIKTVLER 321
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 6
T06874
probable prenyl transferase (EC 2.5.1.-) - Cyanophora paradoxa cyanelle
C:Species: cyanelle Cyanophora paradoxa
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 24-Sep-1999
C:Accession: T06874
R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A.
submitted to the EMBL Data Library, July 1995
A>Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.
A:Reference number: Z15840
A:Accession: T06874
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-323 <STI>
A/Cross-references: EMBL:U0821; NID:g1016083; PIDN:AAA81217.1; PID:g1016130
A:Experimental source: strain Pringsheim LB555
C:Genetics:
A:Gene: preA
A:Genome: cyanelle
C:Superfamily: prenyl transferase A
C:Keywords: cyanelle; transferase

Query Match 30.3%; Score 664.5; DB 2; Length 323;
Best Local Similarity 39.5%; Pred. No. 6e-39;
Matches 147; Conservative 69; Mismatches 97; Indels 59; Gaps 4;

Qy 70 NLVGP---EMSNLTNSIRSLGSGHPISLDTVAKYVVQSSEGHKIRPLMVLMLMAOATEVAPK 126
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 6 NILAPVENELDLTKNLKLVGSCHPILSAASEHLFSASGRPRPAIVLLLSKA-----59
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 127 VQGKEKVVEVPNEGGLAPEVLNDKNPDMMNRSGPLTKDGEIEGTSTNILASORRLAEI 186
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 60 -----TWNEI-----ITSRRHRLAEI 76
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 187 TEMIHAAAILHHDDVIDASETTRNAPSNGAFGNKMMAILAGDFLLGRASVALARLNPEVI 246

Dd	77	TEIHTASTVHDDIIDDESVRKGIPTVHSDPFTKAILAAGDPLFPAQSSWYLANLESLEV	136
Qy	247	ELLATVIANLVGEFMOLKNTVDDAIEATATOETDYILOKTYLKTASIIASCASALL	306
Dd	137	KLISKVITDFAGEGELRR-----GLNPKVDLTVEEYLEKSFYKTASIIAASSPAALL	189
Qy	307	GGATEVEVDAAAAYAGRNLTGAFQIVDMMLDYTSATDLOKPGADILQGLATAPALFAMK	366
Dd	190	SHVDLTVANADLTNYGRHLGLAFQIVDDIIDFSTSEELKPSGSDPLKXGNLTAPALFALE	249
Qy	367	HHAELGPMTKRKFSPOGDVERARELVEKSDGLEKTRALAAEYAOALDIRFPESPAPK	426
Dd	250	ONSELPILOQRPSEPKDPEYTLQIVEETKALEKRELAMEHAQVAIQCLEMLPPSSSKE	309
Qy	427	ALBQLTDKVLTR 438	
Dd	310	ALKLITKVLER 321	

RESULT 7
 AH1818
 solaneseyl diphosphate synthase [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AH1818
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Iriyuchihara, T.;
 Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AH1818
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-323 <KUB>
 A:Cross-references: GB:BA000019; PIDN:BA877620.1; PID:gl7135074; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: sds
 C:Superfamily: prenyl transferase A

```

Query Match 30.2% Score 661.5 DB 2 Length 323;
Best Local Similarity 38.4% Pred. No. 9.7e-39;
Matches 145; Conservative 72; Mismatches 102; Indels 59; Gaps 4;

QY 61 PNOPLINPLNLGPEMSNLTNSIRSLGSGHPSLDITVAKYVYQSEGKHIRPLMLVLLMQA 120
Db 3 PATSLFTPL---VEADLRILADNLKQVLGNRHPILPAAAEHLFGAGGKRRIRPAIVLLISRA 59

QY 121 TEVAKVGQGMKEKVEVPVNNEGIAPEPEVLNDKNPDMNNRSGPLTVDGELGQTSIIASQ 180
Db 60 T-----MDQGITP-----RH 70

QY 181 RLIAETEMIIHAASLLHDVVIDASETRRNAPSGNOAFGNKAILLADGDFLLGSAVALARL 240
Db 71 RLIAETEMIIHAASLLHDVVIDASETRRNAPSGNOAFGNKAILLADGDFLLGSAVALARL 130

QY 241 RNPEVIELLATVIANLVESEFMQLKNVTDDAIEATATQETFPDYVYQKYTLKTASIIAKSC 300
Db 131 DNLQVKKLLSEIYMDLATEIOQGNIRFPASI-----SIEITYEKSYYKTAASIVANSS 183

QY 301 RNSALLGATPEVADAAYVYGRNLGLAFQIVDDMLDYTVSATDLKPKPGADQLQLATAP 360
Db 184 KQAGLSEVSPPTAEHLVYAGRHGLAFQIVDDIIDFISTTTLTKPKVSGDLKSGNLTAP 243

QY 361 ALFAWKHHAELGPMI KKRKTSFGDVERARELVEKSDGLEKTRALAEYAQKALDAIRTFP 420
Db 244 VLFALAEKPYLEVLIEREPFAQEGDLEQALELIIQDSGIIQSSRELAHHHTKLAIEHMLATLP 303

QY 421 ESPARKALEQLTDKVLTR 438
Db 304 PSEHQALIKIAEVAISR 321

```

RESULT 8
573189
Prenyl transferase A - red alga (*Porphyrura purpurea*) chloroplast
C|Species: Chloroplast Porphyrura purpurea
C|Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 24-Sep-1999
C|Accession: 573189
R|Reith, M.; Munholland, J.
Plant Mol. Biol. Rep. 13, 333-335, 1995
A|Title: Complete nucleotide sequence of the *Porphyrura purpurea* chloroplast genome.
A|Reference number: 573108
A|Accession: 573189
A|Status: preliminary; nucleic acid sequence not shown; translation not shown
A|Molecule type: DNA
A|Cross-references: EMBL:U38804; NID:g1276652; PIDN:AAC08154.1; PID:g1276734
A|Note: The nucleotide sequence was submitted to the EMBL Data Library, October 1995
C|Genetics:
A|Gene: preA
A|Genome: chloroplast
C|Superfamily: prenyl transferase A
C|Keywords: chloroplast

Query Match 27.6%; Score 604.5; DB 2; Length 323;
Best Local Similarity 36.1%; Pred. No.9.3e-35;
Matches 135; Conservative 69; Mismatches 111; Indels 59; Gaps 3;

OY 65 LINPLNVGPMSNLTSNIRSLGSHPSLDTVAKYYVQSEGNHRLPLMLMAQAETVA 124
Db LPHPIE---KELYSVEHNLSKVAGTRHPILYAANKHLFEAGKRRLRPALVLVAKST--- 60
OY 125 PKVOGMENVVEVPVNEGFLAPPEVLNDKNPDMMNRSGPLTKOGIEGQTSNNIILASGRLLA 184
Db 61 -----SQQEIKPQGRRLLA 74
OY 185 EITETIHAASSLHDVDIVDASETRRNAPSNGOAFGNKKAILAGPFLIGRAVALARLNPE 244
Db 75 EITETIHAASSLHDVIDDECGSTRGEKTGHKLFTNTKAIVLAGPFLPAOSSWYLANIGNLE 134
OY 245 VIELLATVIANLVGEGFMQLKNVTVDATIEATATOEFPDYVLQKTYLKTAISLAKSGRASA 304
Db LVVKYITKVTITDFAGGEIRQGLVHFDPST-----STDYIEKSPFYTTASLVVASCRGA 187
OY 305 LLGGATREPVADAAYAGRNIAGLAFQIVDMIDYVSKATDI GKPRAGADLQGLATAPALFA 364
Db 188 MLNDLNSQMENDLVLGYKHMGLAFQIMDDVDVLDAGSTKSLGKPSGADFNMGNTLPALFLFA 247
OY 365 WKHNAELGPMIKRKRFSPGCVDERARELVEKSDGLEKTRALAEFAOKALAIRTFPPSPA 424
Db 248 LTQEGKLDQLQREFSPDRDISLALFLIKSSGITTKAKDKAKEQVAAALCCLOLPKXSA 307
OY 425 RKALEQLTDKVLTR 438
Db 308 VSLSKELTHFIITR 321

RESULT 9
GB7775
protein C24A11.9 [imported] - Caenorhabditis elegans
C|Species: Caenorhabditis elegans
C|Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C|Accession: GB7775
R|Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A|Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology
A|Reference numbers: A75000; MIDB:39059613; PMID:9851916
A|Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A|Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A|Accession: GB7775
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-277 <STO>

Db 370 M 370

RESULT 12

AB84759

probable trans-prenyltransferase (imported) - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 24-May-2001

C:Accession: AB84759

R:Lin, X.; Kaul, S.; Rounstey, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB84420; MUID:20083487; PMID:10617197

A:Accession: AB84759

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-297 <STO>

A:Cross-references: GB:AE002093; NID:G3128225; PIDN:AA26705.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2934630

A:Map position: 2

C:Superfamily: prenyl transferase A

Query Match 25.6%; Score 561; DB 2; Length 297;
Best Local Similarity 37.5%; Pred. No. 8.9e-32;
Matches 134; Conservative 61; Mismatches 88; Indels 74; Gaps 6;

Qy 92 PSLDTVAKYVQGS--EKGIRPLMLVLLMAQATEVAPKVGMEKVEVPVNEGGLAPPEVLN 149

Db 7 PKASAAEYFEFGKGVQKQFRSTILLMATALNV-----RVEPALI 47

Qy 150 DKXPDMMNMSGPLTKXGEIEGQTSNLLASORRLAEITEMIHAASLLHDDVIDASERRN 209

Db 48 GSESDIV-----TSELVRQGIATITEMIHASLLHDDVIDADRRG 91

Qy 210 APSGNQAFGNKMAILAGDFLLGRASVALARLRNPEVIELLATVIANLEGEFMOQKNTVD 269

Db 92 VGSILNVVNGK-----VVALATVNEHLVNGETMEITS--- 124

Qy 270 DAIEATATQETFDYVLYOKTYLKTASLIKSCRSALIGATPEYADAAYVGNRLGLAFQ 329

Db 125 ----STEQRYSMDYVYKQTYKTSALISNSCKAAVAVLTGTAEVAVLAFFEGNRLGLAFQ 180

Qy 330 IVDDMLDYTSADTLGKPGADLQGLATPALFAWKHAELGPMIKRKSPDGDVERAR 389

Db 181 LIDILDFGTGTSASLGKGSJSDIRHGVITAPILFAWEFFQLEKVVQVEKDPKRVNDIAL 240

Qy 390 ELVEKSGLEKTRALAEVYAKALDAIRTPESP-----ARKALEQTLTKVLRSR 440

Db 241 EYLKSGKIGIQARLELAHEHNLAAALIGSLPTDNEDVKKSRRLALDLTRVITRKK 297

RESULT 13

E65109

occapprenyl-diphosphate synthase (EC 2.5.1.-) - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002

C:Accession: E65109; PV0010

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.; Rose, D.J.; Mau, B.; Shaoh, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: E65109

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-323 <BLAT>

A:Cross-references: GB:AE000399; GB:U00096; NID:G2367201; PIDN:AA26219.1; PID:G1789578;

A:Experimental source: strain K-12, substrain MG1655

R:Choi, Y.L.; Nishida, T.; Kawamukai, M.; Utsuni, R.; Sakai, H.; Komano, T.

J. Bacteriol. 171, 5222-5225, 1989

A:Title: Cloning and sequencing of an Escherichia coli gene, nlp, highly homologous to l

A:Reference number: JV0027; MUID:89359178; PMID:2670911

A:Accession: PV0010

A:Molecule type: DNA

A:Residues: 11,156-157,196-323 <CHO>

A:Cross-references: GB:X68873; NID:G42128; PIDN:CAA48735.1; PID:G388220

C:Genetics:

A:Gene: nlpB

A:Map position: 69 min

C:Superfamily: prenyl transferase A

C:Keywords: transferase

Query Match 20.6%; Score 451.5; DB 2; Length 323;
Best Local Similarity 31.0%; Pred. No. 4.4e-24;
Matches 116; Conservative 52; Mismatches 143; Indels 63; Gaps 4;

Qy 71 LVGPMSNLTNSIRSLDGSHPSLDTVAKYVQSGKHIRPLMLVLLMAQATEVAPKVGW 130

Db 9 LTAQDMAGVNAALIEQLNSDVQLINQLGYIVSGGKIRPMIAVLARAV-----GY 61

Qy 131 EKVEVPVNEGGLAPPEVLNDKNDPMNMRSGPLTKXGEIEGQTSNLLASORRLAEITEMI 190

Db 62 EGNAAVTI-----AALIEFI 76

Qy 191 HAASLLHDDVIDASETRRNAPSNGQAFGNKMAILAGDFLLGRASVALARLRNPEVIELLA 250

Db 77 HTATLLHDDVIDESDMRKGATANAAGNAAVAVLGDPIYTRAFOMMTSLGSLKYLEVM 136

Qy 251 TVIANLVGEFMOQKNTVDALIEATATQETFDYVLYOKTYLKTASLIKSCRSALLLGAT 310

Db 137 EAVNVIAEGEVLQNMNVDPDI-----TEENYRWYVSTALPEAAACSSILAGCT 189

Qy 311 PEVADAAYVGNRLGLAQIIVDDMLDYTSADTLGKPGADLQGLATPALFAWKHAEL 370

Db 190 PEEKGLDQYGRYLGTAQLIDLDLDYNADEQLKAVNGDDLNKSGKPTPLHHMHGTP 249

Qy 371 LGPMIKRKSPGD-----VERARELVEKSGLEKTRALAEVYAKALDAIRTPESPARK 426

Db 250 EQAQMIRTAIEQGNRHLLFVLLEAMNAGSLEWTRQAEAEADKATIALQVLPDTWRE 309

Qy 427 ALBQTLTKVLRSR 440

Db 310 ALIGLAHIAVQDR 323

RESULT 14

B91137

occapprenyl diphosphate synthase (imported) - Escherichia coli (strain O157:H7, substrain

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: B91137

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kunara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: B91137

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-323 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA837489.1; PID:G13363539; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RMD 0509952

C:Genetics:

A:Gene: BCS4066

C:Superfamily: prenyl transferase A

Query Match 20.6%; Score 451.5; DB 2; Length 323;
Best Local Similarity 30.1%; Pred. No. 4.4e-24;
Matches 116; Conservative 51; Mismatches 132; Indels 87; Gaps 5;

Qy 71 LVGPMSNLTNSIRSLDGSHPSLDTVAKYVQSGKHIRPLMLVLLMAQATEVAPKVGW 130

Db 9 LTAQDMAGVNAAILLEQLNSDVQLINQLGYIYVSGGKREIRPMIAVLAARAV-----GY 61
QY 131 EKVEVPVNEGLAPPEVLNDKPNMNMRSGLTKDGEIEGQTSNIIASQRRRLAEITEMI 190
Db 62 EGNHVTI-----AALIEFI 76
QY 191 HAASLLHDDVDIVASSTRNAPSNGNOACGNKMAIILAGDFLLGRASVALARLNPEVIELLA 250
Db 77 HTATLLHDDVVDESMDRGRKATANAAGNAASVLVGDFIYTRAFQMMTSLGSLKVLVMS 136
QY 251 TVIANLVEGEMQLKNTVDDAIEATATQETPDYLYQKTYLKTASLIASCRASALLGGAT 310
Db 137 EAVNVIAEGEVLQLMNVNDPDI-----TEENYMRVIYSKTARLFEEAAQCSGILAGCT 189
QY 311 PEVADAAYAYGRNIGLAFQIVDDMLDYTVSATDLGKPGADLQGLATAPALFAWKHAE 370
Db 190 PEEKGLQDYGRYLGTAFLQIDLDLDYNADGEQLGKNGVDDLNKPTPLLLHAMHH--- 246
QY 371 LGPMIKKFSDPGDVERARELVEKSDG-----LEKTRALAEYYAQKALD 414
Db 247 -----GTPQQAQMIRTAIEQGNRHLKPVLEAMNACGSLWTRQRAEEADKAIA 297
QY 415 AIRTFPSPARKALEQLTDKVLTRSR 440
Db 298 ALQVLPTPWRREALIGLAHIAVQDR 323

RESULT 15
E85982
octaprenyl diphosphate synthase [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: E85982
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85982
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-323 <STO>
A;Cross-references: GB:AE005174; NID:g12517802; PIDN:AGS8321.1; GSPDB:GN00145; UMGP:Z45
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: ispB
C;Superfamily: prenyl transferase A

Query Match 20.6%; Score 451.5; DB 2; Length 323;
Best Local Similarity 30.1%; Pred. No. 4.4e-24;
Matches 116; Conservative 51; Mismatches 132; Indels 87; Gaps 5;

QY 71 LVGPMSNLTNIRSLGSGHPSLDTVAKYVQSEGGKHIRPLMWLLMAQATEVAPKVGW 130
Db 9 LTAQDMAGVNAAILLEQLNSDVQLINQLGYIYVSGGKREIRPMIAVLAARAV-----GY 61
QY 131 EKVEVPVNEGLAPPEVLNDKPNMNMRSGLTKDGEIEGQTSNIIASQRRRLAEITEMI 190
Db 62 EGNHVTI-----AALIEFI 76
QY 191 HAASLLHDDVDIVASSTRNAPSNGNOACGNKMAIILAGDFLLGRASVALARLNPEVIELLA 250
Db 77 HTATLLHDDVVDESMDRGRKATANAAGNAASVLVGDFIYTRAFQMMTSLGSLKVLVMS 136
QY 251 TVIANLVEGEMQLKNTVDDAIEATATQETPDYLYQKTYLKTASLIASCRASALLGGAT 310
Db 137 EAVNVIAEGEVLQLMNVNDPDI-----TEENYMRVIYSKTARLFEEAAQCSGILAGCT 189
QY 311 PEVADAAYAYGRNIGLAFQIVDDMLDYTVSATDLGKPGADLQGLATAPALFAWKHAE 370
Db 190 PEEKGLQDYGRYLGTAFLQIDLDLDYNADGEQLGKNGVDDLNKPTPLLLHAMHH--- 246
QY 371 LGPMIKKFSDPGDVERARELVEKSDG-----LEKTRALAEYYAQKALD 414

Db 247 -----GTPQQAQMIRTAIEQGNRHLKPVLEAMNACGSLWTRQRAEEADKAIA 297
QY 415 AIRTFPSPARKALEQLTDKVLTRSR 440
Db 298 ALQVLPTPWRREALIGLAHIAVQDR 323

Search completed: January 16, 2003, 09:50:00
Job time : 39 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:41:01 ; Search time 23 Seconds

(without alignments)
793.460 Million cell updates/sec

Title: US-09-830-111A-2

Perfect score: 2194

Sequence: 1 MASPALIRISRSRSIASLR.....ESPAKALEQLDKVLTCSR 440

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1023	46.6	473	1	COOL_YEAST
2	910.5	41.5	378	1	DPS_SCHPO
3	676.5	30.8	323	1	PREA_SYNY3
4	664.5	30.3	323	1	PREA_CVAPA
5	604.5	27.6	323	1	PREA_PORPU
6	571	26.0	323	1	PREA_CVACA
7	451.5	20.6	323	1	ISPB_ECOLI
8	445.5	20.3	329	1	ISPB_HAEIN
9	409	18.6	320	1	HEP2_BACST
10	406	18.5	348	1	HEP2_BACSU
11	365.5	16.7	325	1	IDSA_METH
12	361.5	16.5	324	1	IDSA_METH
13	309	14.1	359	1	GGPP_MYCTU
14	299	13.6	327	1	IDSA_METHA
15	292	13.3	332	1	GGPP_SULSO
16	284	12.9	371	1	GGPP_ARATH
17	274	12.5	366	1	GGPP_SINAL
18	265	12.1	369	1	GGPP_CAPAN
19	260	11.9	299	1	ISPA_ECOLI
20	260	11.9	330	1	GGPP_SULAC
21	256	11.7	357	1	GGPP_CATRO
22	247	11.3	297	1	ISPA_BACST
23	245	11.2	300	1	CRTE_CVAPA
24	210.5	9.6	272	1	ISPA_BACSU
25	210	9.6	291	1	ISPA_MICLU
26	206	9.4	307	1	CRTE_ERWHE
27	203.5	9.3	302	1	CRTE_PANAN
28	199.5	9.1	289	1	CRTE_RHOCA
29	198.5	9.0	288	1	CRTE_RHOSH
30	198.5	9.0	418	1	GGPP_GIBFU
31	191.5	8.7	295	1	ISPA_HAEIN
32	190	8.7	332	1	ISPA_BRAJA
33	184	8.4	332	1	ISPA_RHISN

34	178	8.1	428	1	GGPP_NEUCR	P24322 n geranylge
35	177	8.1	262	1	ISPA_AQUAE	O66952 aquifex ae
36	147	6.7	347	1	PPPS_NEUCR	Q92250 neurospora
37	143.5	6.5	300	1	GGPP_MOUSE	O9wtm0 m geranylge
38	140.5	6.4	294	1	GGPP_BOVIN	P56966 b geranylge
39	139.5	6.4	300	1	GGPP_HUMAN	O95749 h geranylge
40	127.5	5.8	347	1	PPPS_GIBFU	Q92235 gibberella
41	121	5.5	739	1	OCT1_CHICK	P15443 gallus gall
42	119.5	5.4	347	1	PPPS_SCHPO	O14230 schizosacch
43	117	5.3	550	1	SYR_MYCTU	Q10609 mycobacteri
44	115.5	5.3	810	1	CLPC_BACSU	P37571 bacillus su
45	114.5	5.2	282	1	ISPA_BUCAI	P57537 buchnera ap

RESULT 1

COOL_YEAST STANDARD; PRT; 473 AA.

AC P18900;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hexaprenyl pyrophosphate synthetase, mitochondrial precursor

DE (EC 2.5.1.-) (HPS).

GN COQ1 OR YBR003W OR YBR0109.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=D273-10B/Al;

RX MEDLINE=90330660; PubMed=2198286;

RA Ashby M.N., Edwards P.A.;

RT "Elucidation of the deficiency in two yeast coenzyme Q mutants.

RT Characterization of the structural gene encoding hexaprenyl

RT pyrophosphate synthetase.";

RL J. Biol. Chem. 265:13157-13164(1990).

RN [2]

RP SEQUENCE OF i-285 FROM N.A.

RC STRAIN=S288C;

RX MEDLINE=94378721; PubMed=8091860;

RA Wolfe K.H., Lohan A.J.E.;

RT "Sequence around the centromere of Saccharomyces cerevisiae

RT chromosome II: similarity of CEN2 to CEN4.";

RL Yeast 10:S41-S46(1994).

RN [3]

RP SEQUENCE OF 259-479 FROM N.A.

RC STRAIN=S288C;

RA Domdey H., Gassenhuber H., Obermaier B., Piravandi E.;

RT Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: ASSEMBLY OF POLYISOPRENOID SIDE CHAINS. THE POLYPRENYL

CC SYNTHETASE OF COENZYME Q BIOSYNTHESIS CATALYZES THE FORMATION FROM

CC ISOPRENYL DIPHOSPHATE OF ALL TRANS-POLYPRENYL PYROPHOSPHATES

CC GENERALLY RANGING IN LENGTH OF BETWEEN 6 AND 10 ISOPRENE UNITS

CC DEPENDING ON THE SPECIES.

CC -!- PATHWAY: Coenzyme Q biosynthesis; first step.

CC -!- SUBCELLULAR LOCATION: MAY BE PERIPHERALLY ASSOCIATED WITH THE

CC INNER MITOCHONDRIAL MEMBRANE.

CC -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; J05547; AAA34686.1; -.

DR EMBL; Z26494; CAA81272.1; -.

DR EMBL; Z35872; CAAB4939.1; -.
DR PIR; A36625; XUBYTP.
DR SGD; S0000207; COQ1.
DR InterPro; IPR000092; Polyprenyl_synth.
DR Pfam; PF00348; polyprenyl_synth; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW Isoprene biosynthesis; Transferase; Transist peptide; Mitochondrion
TRANSIT 1
FT CHAIN ? 473
FT CHAIN ? 473
SQ SEQUENCE 473 AA; 52559 MW; 676191C9DCA18B3A CRC64; HEXAPRENYL PYROPHOSPHATE SYNTHETASE.

Query Match	1023	46.6%	Score	1023	DB 1	Length	473
Best Local Similarity	46.9%	Pred	No. 7.7e-63				
Matches 223	Conservative	70	Mismatches	118	Indels	64	Gaps

```

Oy 25 RTASAPLRLKCTCTSRSSSSMAAASARLVEPDENPLI--NPLNIVGEMSLTN
Db 4 RSGAAHHIKLISRRCRKFSAVALMAASKLVTP-----KILMNPPLISVXEMNTLAN
Oy 83 IRLSLGCHSPDLTVAKYVOSSEKGIHRLPLMLLMQATIVAPKVC-----GMEKENV
Db 60 IVALIGSHCPVLNKTYSYTFEECKKVRPLLVLLLSRLALEIMETERNHLKIDKSDVPD
Oy 137 PV-----NECLAPPEVLNKNDDMMNRSGPLTKGCELEGOTSNILASORL
Db 120 PIYKPSQNLQFQRPASSISPLHLHGIR-LNPLTGPPEPLPEEFPDKORGLLPQRUL
Oy 184 AEITEMIHAASLLHDVDIVASETRRNA PSGNOAFGNKMAILLADPILGRASVALARLP
Db 179 AEIVEMITASLLHDVDVHDSOTRGRPSGNAAFNTKMAVLADPFLGRATVISRLHP
Oy 244 EVIELAVINLVGEMQAKNVDAIAEAT-----
Db 239 EVELMNSNINLVGEMQAKNKSIDADIDTLENGHKLPLVPSKULEVKEHDFRVPSQ
Oy 276 -----AFOEFPDYVLQTYVLTASTLAKSCRASALIGATPEVADAAYAGRNGL
Db 299 QGLQSHQIIEFAPYVYIHKTYIKTALISKRCALISGASPAVIDECVDFGNLGI
Oy 327 AFOIVDDMLDYTVSATDLCGRAGADLOGLATAPALFAMGHHAELGPMIKRKSFDGVE
Db 359 CFQOLVDDMLDFTVSGKDLGKPSGADLKGATAPVLFAMVEDSLGPLISRNSEKGDVE
Oy 387 RARELVESDQLEKTRALAEYAKALDAIR-FPPSPAKALAEOLTDKVLTSR 440
Db 419 KTIOVRVHNGIATKILAEIRKALQNIQNDSPESDASALEFNLNSILTRRK 473

```

```

RESULT 2
DPS_SCHPO
ID      DPS_SCHPO      STANDARD;      PRT;      378 AA.
AC      043091;
DT      15-JUN-2002 (Rel. 41, Created)
DT      15-JUN-2002 (Rel. 41, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Decaprenyl-diphosphate synthase (EC 2.5.1.-) (Decaprenyl pyrophosphate
DE      synthetase)
DE      DPS_OR SPBP694.01.
OS      Schizosaccharomyces pombe (fission yeast).
OC      Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC      Schizosaccharomycetales; Schizosaccharomycetaceae;
OC      Schizosaccharomycetes.
CX      NCBI_TaxID=4896;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=92729041; Pubmed=9133618;
RA      Suzuki K., Okada K., Kamiya Y., Zhu X.F., Nakagawa T., Kawamukai M.,
RA      Matsuda H.;
RT      "Analysis of the decaprenyl diphosphate synthase (dps) gene in
RT      fission yeast suggests a role of ubiquinone as an antioxidant.";
RL      J. Biochem. 121:496-505(1997).
RN      [2]

```

```

RP SEQUENCE FROM N.A.
RC MEDLINE=21848401; PubMed=11859360;
RX WOOD V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Haldago J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsals K.,
RA James K., Jones L., Jones M., Lecher S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovich E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellern J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voiclaert G., Aert R., Robben J., Grympeier B.,
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaue V., Mottier S.,
RA Gaibler F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rocher M., Galliard C., Tallada V.A., Gazcon A., Thode G.,
RA Dugas R.R., Cuzado L., Jimenez J., Sanchez M., del Rey F., Bento J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerruti L., Lowe T., McCombie W.R., Pallesen I., Potashkin J.,
RA Shpakovski G.V., Useery D., Bartell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe." ;
RL Nature 415:871-880(2002).
CC -I- FUNCTION: SUPPLES DECAPRENYL DIPHOSPHATE, THE PRECURSOR FOR THE
CC SLIDE CHAIN OF THE ISOPRENOID QUINONES UBIQUINONE-10.
CC -I- SIMILARITY: BELONGS TO THE PPP/GSP SYNTHETASIS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; DB4311; BAA12314.1; -.
DR EMBL; DB4311; CAB66154.1; -.
DR InterPro; IPR000092; PolyPrenyl synt.
DR Pfam; PF00348; PolyPrenyl synt. 1.
DR PROSITE; PS00723; POLYPRENYL SYNTHET 1; 1.
DR PROSITE; PS00444; POLYPRENYL SYNTHET 2; 1.
KW Isoprene biosynthesis; Ubiquinone biosynthesis; Transferase.
SQ SEQUENCE 378 AA; 42046 MW; 84B3300F84EBD05B CRC64;

```

[illegible]

```

QY 367 HHAELGPMIKRKSDPGDVERARELVEKSDLEKTRALAEYQAOKALDAIRTFPSPARK 426
DB 305 KYPELGAMIVNRNHPNSDIQARSLVECTDAIEQTITWAKYIKKAKSLCLLPDSPARK 364
QY 427 ALRQLTKVLTRSK 440
DB 365 ALFALADKVIITRK 378

RESULT 3
PREA SYN3
ID PREA SYN3 STANDARD; PRT; 323 AA.
AC P7250; P74758; O08042;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Prenyl transferase (EC 2.5.1.-).
GN PREA OR SLR0611.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugiura M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -!- FUNCTION: POSSIBLE ROLE IN SYNTHESIS OF THE NONAPRENYL SIDE CHAIN
CC OF PLASTOQUINONE OR IN SYNTHESIS OF OTHER PRENYL CHAINS SUCH AS
CC UNDEKAPRENYL PYROPHOSPHATE.
CC -!- SIMILARITY: BELONGS TO THE PFP/GGPP SYNTHETASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90899; BAA16579.2; --
CC BML; D9017; BAA16579.2; JOINED.
CC InterPro: IPR000092; Polyprenyl synt.
CC Pfam: PF00348; polyprenyl synt; 1.
CC PROSITE; PS00444; POLYPRENYL SYNTHET_2; 1.
CC PROSITE; PS00723; POLYPRENYL SYNTHET_1; FALSE NEG.
CC Photosynthesis; Isoprene biosynthesis; Transferase; Complete proteome.
CC SEQUENCE . 323 AA; 35725 MW; 42160CE6C165C685 CRC64;

Query Match 30.8%; Score 676.5; DB 1; Length 323;
Best Local Similarity 40.6%; Pred. No. 2.4e-39;
Matches 153; Conservative 63; Mismatches 102; Indels 59; Gaps 4;

QY 65 LINPLNLVGP---EMSNLTNSIRSLGSGHPSLDTVAKYVQSGKHRIPLMLVLLMAQAT 121
DB 1 MISTTSFAPVDQDLRLTLNKLVLGARHPILGAAAEHLFEAGGKVRVPAIVLLVSRAT 60

```

```

QY 122 EVAPKVGQKEKVEVPVNEGLAPPEVLNDKNPDMNMRSPGLTKDGEIGQTSNIIASOR 181
DB 61 -----LDDQ-----ELTARHR 71
QY 182 RLAEITEMIHAASLLHDDVIDASETRRNAPSGNQAEGNMAILAGDFLLGRASVALARLR 241
DB 72 RLAEITEMIHTASLVHDDVDEADLRNVPVNSLFDNRVAVLAGDFLFAQSSWYLANLD 131
QY 242 NPEVIELLATVIANLVEGEFMQLKNVTVDIAEATATQETFDYLYQTYLKTASLIAKSCR 301
DB 132 NLEVVKLLSEVIRDFAEGBEILQSIINRFD-----TDTLETYLEKSYFKTASLIANSK 184
QY 302 ASALLGGATPEVADAAAYAGYRNLGLAFOIVDDMLDVTVSATDLGKPGADLQGLATAPA 361
DB 185 AAGVLSADPRDVCDDHLYEYKHLGLAFQIVDDITPTSPTEVLGKPGAGSOLISGNITAPA 244
QY 362 LFAWKHHAELGPMIKRKSPDGPVERARELVEKSDLEKTRALAEYQAOKALDAIRTFPE 421
DB 245 LFAWKYPLGLKLIEREFAQAGDLEQALELVEGDDGIRRSRELAANQAQLARQHLVLEM 304
QY 422 SPARKALEOLTDKVLTR 438
DB 305 SAPRESLLELVYVLR 321

RESULT 4
PREA CYAPA
ID PREA CYAPA STANDARD; PRT; 323 AA.
AC P31171;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Prenyl transferase (EC 2.5.1.-).
GN PREA.
OS Cyanophora paradoxa.
OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.
OX NCBI_TaxID=2762;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LB555 / Pringsheim;
RX MEDLINE=91268060; PubMed=1711042;
RX Michalowski C.B., Loeffelhardt W., Bohnert H.J.;
RT "An ORF323 with homology to crtE, specifying prephytoene
RT pyrophosphate dehydrogenase, is encoded by cyanelle DNA in the
RT eukaryotic alga Cyanophora paradoxa.";
RL J. Biol. Chem. 266:11866-11870(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=LB555 / Pringsheim;
RX Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,
RX Bryant D.A.;
RT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa.";
RL Plant Mol. Biol. Rep. 13:327-332(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=LB555 / Pringsheim;
RX Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,
RX Farley J.Y., Schluchter W.M., Chung S., Newmann-Spallart C.,
RX Steiner J.W., Jakowitsch J., Bohnert H.J., Bryant D.A.;
RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:
RT the genetic complexity of a primitive plastid.";
RL (in) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,
RL Schwemmler W. (eds.);
RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg
RL (1997).
RN [4]
RP SEQUENCE OF 1-46 FROM N.A.
RX STRAIN=LB555 / Pringsheim;
RX MEDLINE=91117189; PubMed=2126059;
RX Michalowski C.B., Pfanagl B., Loeffelhardt W., Bohnert H.J.;
RT "The cyanelle S10 spc ribosomal protein gene operon from Cyanophora
RT paradoxa.";

```

```

RL Mol. Gen. Genet. 224:222-231(1990).
CC -1- FUNCTION: POSSIBLE ROLE IN SYNTHESIS OF THE NONAPRENYL SIDE CHAIN
CC OF PLASTOQUINONE OR IN SYNTHESIS OF OTHER PRENYL CHAINS SUCH AS
CC UNDEKAPRENYL PYROPHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE PPP/GGPP SYNTHETASES FAMILY.
CC -1- CAUTION: Was originally (Ref.1) called crte.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M37111; AAA65472.1; -
CC EMBL; U30821; AAA81217.1; -
CC EMBL; M30487; AAA63631.1; -
CC PIR; A40433; A40433.
CC InterPro; IPR000092; Polyprenyl_synth.
CC Pfam; PF00348; polyprenyl_synth_1.
CC PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
CC PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
CC Photosynthesis; Isoprene biosynthesis; Transferase; Cyanelle.
SQ SEQUENCE 323 AA; 35919 MW; CB4802466342B09A CRC64;

Query Match 30.3%; Score 664.5; DB 1; Length 323;
Best Local Similarity 39.5%; Pred. No. 1.6e-38;
Matches 147; Conservative 69; Mismatches 97; Indels 59; Gaps 4;

QY 70 NLVGP---EKSNTLSNIRSLGSHPSLDTVAKYVYOSGKHIRPLMLMAQATEVA 126
DB 6 NILAPVENELDLTKNKKLVGSGHPILSAASEHLFSASGRRPPAIVLLISK- 59
QY 127 VQGEKVEVEPVNEGGLAPPEVLNDKNPDMNMRSGPLTKDGEIGQTSNIIASORLAEI 186
DB 60 -----TMENET-----ITSKRRLAEI 76
QY 187 TEMIHAASLLHDDVIDASERRRNPASGNQAFGNKMAIILAGDFLLGRASVALRLRNEVI 246
DB 77 TEIHTASLVHDDILDESDDRGRIPVHSPGTKIALIADGFLPAQSGSYLVANLESLEV 136
QY 247 ELATVIYANLVEGEFMQKNTVDDAIEATATQETFDYIYQKTYLKTASLIASCRASAL 306
DB 137 KLISKVIITDPAEGEIR-----GLNQFVVDLTLEBYLEKSPFKYKTASLIASAASKAALL 189
QY 307 GGATPEVADAAYVAGRNGLAFQIVDDMLDYVSATDLGKPPAGADLQGLATPAPALPAWK 366
DB 190 SHVDLTIVANDLYNGRHGLAFQIVDDILDTSTBELGKPSGSDLLKGNLTAPVLFAL 249
QY 367 HHAELGPMIRKPSDGDVERARELVEKSDGLEKTRALAEYAQKALDAIRTPESPARK 426
DB 250 QNSLILIGLQRFSEPDPEFTYIQIVETKAIKTRRLAMEHAQVVAIQCEINLPSSSKE 309
QY 427 ALBQITDKVLT 438
DB 310 ALKLITKYVLER 321

RESULT 5
PREA_PORPU STANDARD; PRT; 323 AA.
ID PREA_PORPU
AC P51268;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-1999 (Rel. 38, Last annotation update)
DE Prenyl transferase (EC 2.5.1.-).
GN Porphyra purpurea.
OS Porphyra purpurea.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxID=27877;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-Avonport;
RT Ketch M.E., Munholland J.;
RT "Complete nucleotide sequence of the porphyra purpurea chloroplast
RT genome."
RT Plant Mol. Biol. Rep. 13:333-335(1995).
CC -1- FUNCTION: POSSIBLE ROLE IN SYNTHESIS OF THE NONAPRENYL SIDE CHAIN
CC OF PLASTOQUINONE OR IN SYNTHESIS OF OTHER PRENYL CHAINS SUCH AS
CC UNDEKAPRENYL PYROPHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE PPP/GGPP SYNTHETASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U38804; AAC08154.1; -
CC InterPro; IPR000092; Polyprenyl_synth.
CC Pfam; PF00348; polyprenyl_synth_1.
CC PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
CC PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
CC Photosynthesis; Isoprene biosynthesis; Transferase; Chloroplast.
SQ SEQUENCE 323 AA; 35589 MW; 9FB02FE42A2019B2 CRC64;

Query Match 27.6%; Score 604.5; DB 1; Length 323;
Best Local Similarity 36.1%; Pred. No. 2.1e-34;
Matches 135; Conservative 69; Mismatches 111; Indels 59; Gaps 3;

QY 65 LNPVLNVPKSNLTSNIRSLGSHPSLDTVAKYVYOSGKHIRPLMLMAQATEVA 124
DB 7 LPHPIE---KELSYVENHLKSVAGTRHPILYAANKHLEAGCKRLRPAIVLVAKST--- 60
QY 125 PVQGEKVEVEPVNEGGLAPPEVLNDKNPDMNMRSGPLTKDGEIGQTSNIIASORLAEI 184
DB 61 -----SQGEIRKQGRRLA 74
QY 185 EYTEMIHAASLLHDDVIDASERRRNPASGNQAFGNKMAIILAGDFLLGRASVALRLRNEVI 244
DB 75 ETEIHTASLVHDDILDESDDRGRIPVHSPGTKIALIADGFLPAQSGSYLVANLESLEV 134
QY 245 VIELATVIYANLVEGEFMQKNTVDDAIEATATQETFDYIYQKTYLKTASLIASCRASA 304
DB 135 VYKVIITKYITDPAEGEIRQGLVHFPESI-----SIDDYIEKSFYKTASLIASAASCGAA 187
QY 305 ILGGATPEVADAAYVAGRNGLAFQIVDDMLDYVSATDLGKPPAGADLQGLATPAPALPA 364
DB 188 MUNDINSQWMDLVLYGKRMGLAFQIMDDVDLDIAGSTKSLGKPPGADPMGNLTAPILFA 247
QY 365 WGHAEELGPMIRKPSDGDVERARELVEKSDGLEKTRALAEYAQKALDAIRTPESPARK 424
DB 248 LTQEGKLDQGLQRFESDSDRDISLALFLIKSGGITKAKDLAKEQVAAALCCLQFLPNSAP 307
QY 425 KRALBQITDKVLT 438
DB 308 VSLKELTHFIITR 321

RESULT 6
PREA_CYACA STANDARD; PRT; 323 AA.
ID PREA_CYACA
AC Q9TSL1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Prenyl transferase (EC 2.5.1.-).
GN Cyanidium caldarium.
OS Cyanidium caldarium.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OX Cyanidium.
RN [1]

```

```

OX NCBI_TaxID=2771;
RN SEQUENCE FROM N.A.
RC STRAIN=K-1;
RX MEDLINE=2049659; PubMed=11040290;
RA Gloeckner G., Rosenthal A., Valentin K.-U.;
RT "The structure and gene repertoire of an ancient red algal plastid
RT genome.";
RL J. Mol. Evol. 51:382-390(2000).
CC -!- FUNCTION: POSSIBLE ROLE IN SYNTHESIS OF THE NONAPRENYL SIDE CHAIN
CC OF PLASTOQUINONE OR IN SYNTHESIS OF OTHER PRENYL CHAINS SUCH AS
CC UNDEKAPRENYL PYROPHOSPHATE.
CC -!- SIMILARITY: BELONGS TO THE PPP/GGPP SYNTHETASES FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC DR EMBL; AF021186; AAF12896.1; -
CC DR InterPro; IPR000092; Polyprenyl synt.
CC DR Pfam; PF00348; polyprenyl synt.; 1
CC DR PROSITE; PS00723; POLYPRENYL SYNTHET 1; 1.
CC DR PROSITE; PS00444; POLYPRENYL SYNTHET 2; 1.
CC DR Photosynthesis; Isoprene biosynthesis; Transferase; Chloroplast.
CC SQ SEQUENCE 323 AA; 36052 MW; C9E7F23CBED3FC47 CRC64;
CC
CC
CC Query Match 26.0%; Score 571; DB 1; Length 323;
CC Best Local Similarity 35.8%; Pred. No. 4.1e-32;
CC Matches 133; Conservative 67; Mismatches 110; Indels 64; Gaps 4;
CC
CC QY 69 LNLVGPMSNLTSNIRSLGSGHPSLDTVAKYVQSGKIRPLMWLLMAQATEVAPKVQ 128
CC Db 8 LQSVKEDLLNIEQTLNKLKVNPNILSAAXHLLVSKKIRPAIVLLVAKAI----- 60
CC
CC QY 129 GWKEVVPVNEGLAPPEVLNDKNPDMNMRSGPLTKDGEIGQTSNILASQRLAEITE 188
CC Db 61 -----DKN-----KKIKTSQQRLEAVTE 78
CC
CC QY 189 MIHAASLLHDDVIDASETRNAPSGNAGFKGMALAGDFLGPASVALARLNPEVIEL 248
CC Db 79 IHTATLLHDDVDSEIRRGTSKWNTPGKNIAVFAFGDFLFAQSSWYLANINNLVYKA 138
CC
CC QY 249 LATVIANLVEGEFQNLKNTVDDEATATQTPDY-----YLOKTYLKTASLIKSCRSA 304
CC Db 139 ISKVITDLAEGELQ-----NLQNTYYSIKVLEKSFNTASLIASCKSCC 187
CC
CC QY 305 LGGATEVADAAYAYGRNLGAFQIVDDMLDYTVSATDLGPKAGADLQGLATAPALFA 364
CC Db 188 LLSDFDQSLNSKFNFGKLNGLAFQIIDDITSSSTALGKMTTSDLGLNLTAPVLA 247
CC
CC QY 365 WKHAEIGPMIKRFPDGVVERRELVEKSDGLEKTRALAEVAQAQALDAIRTFPSPA 424
CC Db 248 LTKNSKLFKIEREFCKSDISEAINIKETNAIESFSLAYEHIEAINSIKDLPTSSE 307
CC
CC QY 425 KALEQLTQDKVLTR 438
CC Db 308 KDSLIEIAYDLLNR 321
CC
CC RESULT 7
CC ISPB_ECOLI
CC ID ISPB_ECOLI STANDARD; PRT; 323 AA.
CC AC P19641;
CC DT 01-FEB-1991 (Rel. 17, Created)
CC DT 01-FEB-1994 (Rel. 28, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Octaprenyl-diphosphate synthase (EC 2.5.1.-) (Octaprenyl pyrophosphate
CC synthetase) (OPP synthetase).
CC GN ISPB OR CEL OR B3187.

```

```

OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RX MEDLINE=94146411; PubMed=8312607;
RA Jeong J.H., Kitakawa M.S., Isono S., Isono K.;
RT "Cloning and nucleotide sequencing of the genes, rpiu and rpmA, for
RT ribosomal proteins L21 and L27 of Escherichia coli.";
RL DNA Seq. 4:59-67(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE OF 196-323 FROM N.A.
RX MEDLINE=89359178; PubMed=2670911;
RA Choi Y.-L., Nishida T., Kawamukai M., Utsumi R., Sakai H., Komano T.;
RT "Cloning and sequencing of an Escherichia coli gene, nlp, highly
RT homologous to the ner genes of bacteriophages Mu and D108.";
RL J. Bacteriol. 171:5222-5225(1989).
RN [4]
RP FUNCTION.
RC STRAIN=K12 / JM109;
RX MEDLINE=94311902; PubMed=8037730;
RA Asai K.-I., Fujisaki S., Nishimura Y., Nishino T., Okada K.,
RA Nakagawa T., Kawamukai M., Matsuda H.;
RT "The identification of Escherichia coli ispB (cel) gene encoding the
RT octaprenyl diphosphate synthase.";
RL Biochem. Biophys. Res. Commun. 202:340-345(1994).
CC -!- FUNCTION: SUPPLIES OCTAPRENYL DIPHOSPHATE, THE PRECURSOR FOR THE
CC SIDE CHAIN OF THE ISOPRENOID QUINONES UBIQUINONE AND MENAQUINONE.
CC -!- SIMILARITY: BELONGS TO THE PPP/GGPP SYNTHETASES FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC DR EMBL; D13267; -; NOT ANNOTATED_CDS.
CC DR EMBL; U18997; AAA57988.1; -
CC DR EMBL; A5000399; AAC76219.1; -
CC DR EMBL; X68873; CAA48735.1; ALT_SEQ.
CC DR PIR; PV0010; PV0010.
CC DR EcoGene; EG10017; ispB.
CC DR InterPro; IPR000092; Polyprenyl synt.
CC DR Pfam; PF00348; polyprenyl synt.; 1.
CC DR PROSITE; PS00444; POLYPRENYL SYNTHET 2; 1.
CC DR PROSITE; PS00723; POLYPRENYL SYNTHET 1; 1.
CC DR Isoprene biosynthesis; Transferase; Complete proteome.
CC SQ SEQUENCE 323 AA; 35217 MW; 08AD7AE7AC230EF8 CRC64;
CC
CC
CC Query Match 20.6%; Score 451.5; DB 1; Length 323;
CC Best Local Similarity 31.0%; Pred. No. 6.3e-24;
CC Matches 116; Conservative 52; Mismatches 143; Indels 63; Gaps 4;
CC
CC QY 71 LVGPMSNLTSNIRSLGSGHPSLDTVAKYVQSGKIRPLMWLLMAQATEVAPKVQGW 130
CC Db 9 LTAQDMAGVNAALIEQLNSDVQLINQLNGYIVSGGKIRIEMIAVLAARAV-----GY 61
CC
CC QY 131 EKVEVPVNEGLAPPEVLNDKNPDMNMRSGPLTKDGEIGQTSNILASORRLAEITE 190
CC Db 131 EKVEVPVNEGLAPPEVLNDKNPDMNMRSGPLTKDGEIGQTSNILASORRLAEITE 190

```

```

Db      62 EGNAAHTI-----ALIEFI 76
QY      191 HAASLLHDDVIDASETRRNAPSGNQAFGNKAILAGDFLLGRASVALARLNPEVIELLA 250
       77 HTATLLHDDVDVDESDMRGRATANAAGNAAASYLVGFITRAFOVMTSISGSIKVLEWVS 136
QY      251 TVIANLVEGEFMQKNTVDIAIEATATQETFDYLOKTYIKTSLAKSCRASALLGGAT 310
       137 EAVNVIAEGEVLQDMNVNDPDI-----TEENVRYIYKTKARLFEAAQCGIILAGCT 189
QY      311 PEVADAAYAGRNGLAFOIYVDDMLDYTVSATDLGKPGADLOGLATAPALPAMKHAE 370
       190 PEEKGLQDGRYIGTAFOLLIDLDLDYNADGEOGKNGVGDINEGKPTLPILHAMHGT 249
QY      371 LGMPIKRFSDPGD---VERARELVKSDGLEKTRALAEVYOKALDAIRTPESPARK 426
       250 EOKMIRTAIEQGNRHLLEPVLNANACGSLLEWTRORAEEDAKAIALQVLPDTPWRE 309
QY      427 ALQOLDKVLTFRSR 440
Db      310 ALIGLAHIAVORDR 323

RESULT 8
ID      ISPB_HAEIN      STANDARD;      PRT;      329 AA.
AC      P44916;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Octaprenyl-diphosphate synthase (EC 2.5.1.-) (Octaprenyl pyrophosphate
       synthetase) (OPP synthetase).
GN      ISPB OR HI0881.
OS      Haemophilus influenzae.
OC      Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
       Haemophilus.
OX      NCBI_TaxID=727;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=Rd / KM20 / ATCC 51907;
       MEDLINE=95350630; PubMed=7542800;
       Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
       Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
       McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
       Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
       Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
       Utechtack T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
       Fine L.D., Frichman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
       Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
       Venter J.C.;
       "Whole-genome random sequencing and assembly of Haemophilus influenzae
       Rd.";
       Science 269:496-512 (1995).
RL      [2]
RP      IDENTIFICATION BY MASS SPECTROMETRY.
RX      MEDLINE=20137488; PubMed=10675023;
       Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
       Gray C., Fountoulakis M.;
       "Two-dimensional map of the proteome of Haemophilus influenzae";
       Electrophoresis 21:411-429 (2000).
RT      FUNCTION: SUPPLIES OCTAPRENYL DIPHOSPHATE, THE PRECURSOR FOR THE
       SIDE CHAIN OF THE ISOPRENOID QUINONES UBIQUINONE AND MENAQUINONE
       (BY SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO THE PPP/GPPP SYNTHETASES FAMILY.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).

```

```

DR      EMBL: U32770; AAC22540.1; -.
DR      TIGR: HI0881; -.
DR      InterPro: IPR000092; Polyprenyl_synth.
DR      Pfam: PF00348; Polyprenyl_synth; 1.
DR      PROSITE: PS00444; POLYPRENYL_SYNTHET_2; 1.
DR      PROSITE: PS00723; POLYPRENYL_SYNTHET_1; 1.
KW      Isoprene biosynthesis; Transferrase; Complete proteome.
SQ      SEQUENCE 329 AA; 35911 MW; 65DBC4CCDAD072B04 CRC64;

Query Match      20.3%; Score 445.5; DB 1; Length 329;
Best Local Similarity 32.5%; Pred. No. 1,7e-23;
Matches 119; Conservative 49; Mismatches 133; Indels 65; Gaps 6;

QY      71 LVGPEMSNLTSIRSLGSGHPSLDYVAKYVYSGEGRIRLMLWMAQAEVAPKQGW 130
       15 LADPMQKVNQNIILQINSDVPLIGQLGFYIVQGGKRIPLVILARSL----- 65
QY      131 EKVEVPVNEGLAPPEVLNDKNPDMNMNRSGLPTKDGIEGTSNIIASQRLAITEMI 190
       66 -----GFESNSITCNT-----PVEFI 82
QY      191 HAASLLHDDVIDASETRRNAPSGNQAFGNKAILAGDFLLGRASVALARLNPEVIELLA 250
       83 HTASLLHDDVDVDESDMRGRATANAAGNAAASYLVGFITRAFOVLAQLESILKILSMA 142
QY      251 TVIANLVEGEFMQKNTVD-DAIENTATQETFDYLOKTYIKTSLAKSCRASALLGGA 309
       143 DATNVLAEGEVQQLNVNDPDTSEAN-----YMRVYKTKARLFEVAAQAAIYVAGG 194
QY      310 PEVADAAYAGRNGLAFOIYVDDMLDYTVSATDLGKPGADLOGLATAPALPAMKH-H 368
       195 TTAQEKALQDYGRIYIGTAFOLLVDVLDYSANTQALGKNGVGDINEGKPTLPILHAMHGN 254
QY      369 AELGMPIKRFSDPGDVERARE---LVKESDLEKTRALAEVYOKALDAIRTPESPARK 425
       255 AQQAALIREALIEQGGKRAIDEVALIMTEHKSIDYANMRKAKEAQAQVDAIEILPESEYK 314
QY      426 KALBOL 431
Db      315 QALISL 320

RESULT 9
ID      HEP2_BACST      STANDARD;      PRT;      320 AA.
AC      P55785;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Hepcaprenyl diphosphate synthase component II (EC 2.5.1.30) (HEPPP
       synthase subunit 2).
GN      HEP2 OR HEPS-2.
OS      Bacillus stearothermophilus.
OC      Bacteria; Firmicutes; Bacillales; Geobacillus.
OX      NCBI_TaxID=1422;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=ATCC 10149;
       MEDLINE=95355463; PubMed=7629164;
       Kojima-Takeshita A., Koyama T., Obara S., Ogura K.;
       "Molecular cloning and nucleotide sequences of the genes for two
       essential proteins constituting a novel enzyme system for hepcaprenyl
       diphosphate synthesis";
       J. Biol. Chem. 270:18396-18400 (1995).
RL      [2]
RP      FUNCTION: SUPPLIES HEPCAPRENYL DIPHOSPHATE, THE PRECURSOR FOR THE
       SIDE CHAIN OF THE ISOPRENOID QUINONE MENAQUINONE-7 (MQ-7).
CC      -1- CATALYTIC ACTIVITY: All-trans-hexaprenyl diphosphate + isopentenyl
CC      diphosphate = diphosphate + all-trans-heptaprenyl diphosphate.
CC      -1- SUBUNIT: HETERODIMER OF COMPONENT I AND II.
CC      -1- SIMILARITY: BELONGS TO THE PPP/GPPP SYNTHETASES FAMILY.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D49976; BAA08725.1; --
CC InterPro; IPR000092; Polyprenyl_synth.
CC Pfam; PF00348; polyprenyl_synth; 1.
CC PROSITE; PS00444; POLYPRENYL SYNTHET 2; 1.
CC PROSITE; PS00723; POLYPRENYL SYNTHET 1; 1.
CC Isoprene biosynthesis; Transferase.
CC SEQUENCE 320 AA; 35808 MW; A51B6630E6B88F46 CRC64;

Query Match 18.6%; Score 409; DB 1; Length 320;
Best Local Similarity 28.5%; Pred. No. 5e-21;
Matches 106; Conservative 61; Mismatches 127; Indels 78; Gaps 6;

QY 70 NLVGPENSLTNSIRLLSGHPSLDTVAKYVQSEKHIRPLMLVILMAQATEVAPKVQ 129

Db 8 SFUSDDLAAVEEELERAVQSEYGLGEALHLLQAGGKRIRPVFVLLAARF----- 58

QY 130 WEKVVEVPVNEGLAPPEVLNDKNDPMNMRSGPLTKDGEIETGQTSNILASORRLAEITEM 189

Db 59 -----GQYD--LERMKHVAVALEL 75

QY 190 IHAASLLHDDVIDASSTRNAPSGNQAFGNKMAILAGDFLLGRASVALARLNPEVIELL 249

Db 76 IHMASLVHDDVIDDADLRGRPTIKAKWSNFAMVTGDFLFAFSLERMAELGNPRAHQLV 135

QY 250 ATVIANLVEGEFMQLKNTVDDAIEATATQETFDY----YLOKTYLKTASLIASCRASAL 305

Db 136 AKTIVEVCRGEIEQIKD-----KYRFDQFLRYLRIRRKRTALLIAASCQLGAL 184

QY 306 LGGATPEVADAAYAGRNIGLAFQIVDDMLDVTYSATDLGKPGADLQGLATAPALFAW 365

Db 185 AAGAPEIVKRLVWFGVHGVMSFOITDILFTGTETQOLGKPGAGSDLLQGNVTLPVLIAL 244

QY 366 KKH-----AELGPMIKRKFPDGDVERARELVKSGLEKTRALABEYAKALDAIRTF 419

Db 245 SDERVKAIAAVGPE-----TDVAEMAIVSAIKRTDAIERSYALSRYLDKALHLLDGL 299

QY 420 PESPARKALEQL 431

Db 300 PMNEARGLLRDL 311

RESULT 10
HEP2_BACSU STANDARD; PRT; 348 AA.
ID HEP2_BACSU
AC P31114;
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Heptaprenyl diphosphate synthase component II (EC 2.5.1.30) (HEPPP
DE synthase subunit 2) (Spore germination protein C3).
GN HEP2 OR HEPB OR GERC OR GERC3.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]

SEQUENCE FROM N.A.
RC STRAIN=168;
RA Henner D.J.;
RT "Sequence of Bacillus subtilis dbpA, mtr(A,B), gerC(1-3), ndk, cheR,
RT aro(B,E,F,H), trp(A-F), hisH, and tyrA genes.";
RL Submitted (JAN-1992) to the EMBL/GenBank/DBJ databases.
RN [2]

CHARACTERIZATION OF GERC LOCUS.
RX MEDLINE=91037938; PubMed=2121900;
RA Yazdi M.A.; Moir A.;
RT "Characterization and cloning of the gerC locus of Bacillus subtilis
RT 168.";

RL J. Gen. Microbiol. 136:1335-1342(1990).
RN [3]
RP FUNCTION.

RX MEDLINE=98386502; PubMed=9720033;
RA Leatherbarrow A.J.H., Yazdi M.A., Curson J.P., Moir A.;

RT "The gerC locus of Bacillus subtilis, required for menaquinone
RT biosynthesis, is concerned only indirectly with spore germination.";

RL Microbiology 144:2125-2130(1998).

CC -!- FUNCTION: SUPPLIES HEPTAPRENYL DIPHOSPHATE, THE PRECURSOR FOR THE
CC SIDE CHAIN OF THE ISOPRENOID QUINONE MENAQUINONE-7 (MQ-7).

CC -!- CATALYTIC ACTIVITY: All-trans-hexaprenyl diphosphate + isopentenyl
CC diphosphate = diphosphate + all-trans-heptaprenyl diphosphate.

CC -!- SUBUNIT: HETERODIMER OF COMPONENT I AND II.

CC -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M80245; AAA20856.1; --
DR EMBL; Z99115; CABI4190.1; --

DR Subtilist; BG10281; hept.

DR InterPro; IPR000092; Polyprenyl_synth.

DR Pfam; PF00348; polyprenyl_synth; 1.

DR PROSITE; PS00444; POLYPRENYL SYNTHET 2; 1.

DR PROSITE; PS00723; POLYPRENYL SYNTHET 1; 1.

DR Sporulation; Isoprene biosynthesis; Transferase; Complete proteome.

KW SEQUENCE 348 AA; 39515 MW; OFF9C9199F04BE CRC64;

SQ

Query Match 18.5%; Score 406; DB 1; Length 348;
Best Local Similarity 27.7%; Pred. No. 9.1e-21;
Matches 101; Conservative 69; Mismatches 125; Indels 70; Gaps 6;

QY 70 NLVGPENSLTNSIRLLSGHPSLDTVAKYVQSEKHIRPLMLVILMAQATEVAPKVQ 129

Db 36 SFLNDDIDVIERELEQTVRSYDPLSEAGLHLLQAGGKRIRPVFVLL----- 82

QY 130 WEKVVEVPVNEGLAPPEVLNDKNDPMNMRSGPLTKDGEIETGQTSNILASORRLAEITEM 189

Db 83 -----SGMF--GDYD-----INIKYVAVTLEM 103

QY 190 IHAASLLHDDVIDASSTRNAPSGNQAFGNKMAILAGDFLLGRASVALARLNPEVIELL 249

Db 104 IHMASLVHDDVIDDADLRGRPTIKAKWSNFAMVTGDFLFAFSLERMAELGNPRAHQLV 163

QY 250 ATVIANLVEGEFMQLKNTVDDAIEATATQETFDYLYLOKTYLKTASLIASCRASALLGGA 309

Db 164 SQTIVEVCLGEIEQIKDKYN-----MEQNLRTYLRIRKRKTALLIAVSCQLGAIASGA 216

QY 310 TPEVADAAYAGRNIGLAFQIVDDMLDVTYSATDLGKPGADLQGLATAPALFAWKHHA 369

Db 217 DEKHKALYVFGYVGVMSYQIIDDILFTSTEEELGKPGVGGDLLQGNVTLPVLIALKN-- 274

QY 370 ELGPMIKRKE-----SDPGDVERARELVKSGLEKTRALABEYAKALDAIRTFPESP 423

Db 275 ---PALKNQKLINSETTQQLPEIIEIKKTAIEASMAVSEMYLQKAFQKLTLPGR 331

QY 424 ARKAL 428

Db 332 ARSSL 336

RESULT 11
IDSA_METHH STANDARD; PRT; 325 AA.

AC O26156;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)


```

DE Bifunctional short chain isoprenyl diphosphate synthase [includes:
DE Farnesyl pyrophosphate synthetase (EC 2.5.1.1) (FPP synthetase)]
DE (Dimethylallyltransferase); Geranyltransferase (EC 2.5.1.10)].
GN IDSA OR MTH50.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN 11)
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; Pubmed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Delongchery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzaeh R., Blakey D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pochter B., Qiu D.,
RA Spadofora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA Duggan J., Shiner G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-T., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
CC diphosphate = diphosphate + geranyl diphosphate.
CC -1- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC = diphosphate + trans,trans-farnesyl diphosphate.
CC -1- COFACTOR: REQUIRES MAGNESIUM.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB000797; AAB84557.1; -
CC InterPro: IPR000092; Polyprenyl_synth.
CC Pfam: PF00348; Polyprenyl_synth: 1.
CC PROSITE: PS00723; POLYPRENYL_SYNTHET_1; 1.
CC PROSITE: PS00444; POLYPRENYL_SYNTHET_2; 1.
CC Lipid synthesis; Isoprene biosynthesis; Transferase; Magnesium;
CC Multifunctional enzyme; Complete proteome.
CC KW
CC SEQUENCE 325 AA; 35485 MW; BCF3A285A207916E CRC64;
SQ
Query Match 16.7%; Score 365.5; DB 1; Length 325;
Best Local Similarity 29.3%; Pred. No. 4,9e-18;
Matches 109; Conservative 61; Mismatches 133; Indels 69; Gaps 9;
OY 75 EMSLTSNIRSLGSGHP-SLDTVAKYVQSEGHIRPLMLWMAQATEVAPKVGMEKV 133
DB 13 EMD--ERIESISDITPERILLRASEHLITAGGKIRPSLALLSSEAVGDP----- 62
OY 134 VEVVNEGLAPREVLNDKNPDMNMRSGPLTKDGEIEGQTSNIIASQRLLAETEMIHAA 193
DB 63 -----GDAAG-----VAAIETLHTF 78
OY 194 SLHDDVIDASETRRNAPSGNQAFGKMAIIAGDFLLGRASVALARKRNEVL-ELLATV 252
DB 79 SLHDDIMDDDEIRGEPRAVHVLGEPMAIIAGDVLFSKAFVIRNGSEWKEALAVV 138
OY 253 IANLV---EGEFMOLKNTVDATATQENFPDYLLQTKVLTASLAKSCRAISALLGCA 309
DB 139 VDSCKIKCEGQALDM--GFEERLDVTEE-----YMETIKTKTALLAAATKAGALWGG 191
OY 310 TPEVADAAYVYGRNLGLAFOIVDDMLDYVSATDLGKRPAGADQLGLATAPALFAWKHNA 369
DB 192 SPOEIALBEDYGCIGLAFQIHDDYLDVDSDESLGKRPVSDIAEGKMTLMVVKALERAS 251
OY 370 ELGPMIKRKSDFGD---VERARELVEKSGLEKTRALAEYAKALDAIRTPESPARK 426

```

```

DB 252 EKDRERLSILGSGEKLVAERIEFERYGATEYAHAVLDHVMARELVEESDARE 311
OY 427 ALEOLDKVLTR 438
DB 312 ALAMTADPFLER 323
RESULT 12
IDSA_METTM STANDARD; PRT; 324 AA.
AC 053479;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bifunctional short chain isoprenyl diphosphate synthase [includes:
DE Farnesyl pyrophosphate synthetase (EC 2.5.1.1) (FPP synthetase)]
DE (Dimethylallyltransferase); Geranyltransferase (EC 2.5.1.10)].
GN IDSA.
OS Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).
OC Archaea; Euryarchaeota; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=79929;
RN 11)
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=95070145; Pubmed=7979381;
RA Chen A., Poulter C.D.;
RT "Isolation and characterization of idsa: the gene for the short chain
RT isoprenyl diphosphate synthase from Methanobacterium
RT thermoautotrophicum.";
RL Arch. Biochem. Biophys. 314:399-404(1994).
CC -1- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
CC diphosphate = diphosphate + geranyl diphosphate.
CC -1- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC = diphosphate + trans,trans-farnesyl diphosphate.
CC -1- COFACTOR: REQUIRES MAGNESIUM.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S75695; AAB32421.1; -
CC InterPro: IPR000092; Polyprenyl_synth.
CC Pfam: PF00348; Polyprenyl_synth: 1.
CC PROSITE: PS00723; POLYPRENYL_SYNTHET_1; 1.
CC PROSITE: PS00444; POLYPRENYL_SYNTHET_2; 1.
CC Lipid synthesis; Isoprene biosynthesis; Transferase; Magnesium;
CC Multifunctional enzyme.
CC KW
CC INTITMET 0
CC SEQUENCE 324 AA; 35505 MW; 88558A8ADDF5F66C1 CRC64;
SQ
Query Match 16.5%; Score 361.5; DB 1; Length 324;
Best Local Similarity 31.4%; Pred. No. 9,2e-18;
Matches 107; Conservative 58; Mismatches 134; Indels 42; Gaps 8;
OY 121 TEVAPKVGMEKVEVPVNEGLAPREVLNDKNPDMNMRSGPLTKDGEIEGQTSNIIAS- 179
DB 1 TEVLILRKYSVADKRI-----MECISDITPPTLLKASHLITAGGKIRPSLALLSC 54
OY 180 -----QRLAETEMITAAASLHDVDVDASETRNAPSGNQAFGKMAIIAGDFLL 230
DB 55 EAVGNPEADAGVAAAIETLHTFSLIHDDIMDDEMRGEPSVAVINGEPMAIIAGDVL 114
OY 231 GRASVALARKRNP-VIELATVIANLV---EGEFMOLKNTVDATATQENFPDYLLQ 286
DB 115 SKAFPAVIRNGDSERVKALAVVDSCKIKCEGQALDM--GFEERLDVTE-----DYM 167

```


QY 287 KTYLKTASLAKSCRASALGGATPEVADAAAYAGRNGLAFQIVDDMLDYTSATDLGK 346
 DB 168 MIYKTAALIAAATKAAGIMGGASEREVEALDYGKFIQGLAFQIHDYLDVDSDESLGK 227
 QY 347 PAGADLQGLGATAPALFAWKHAE-----LGPIMKKFSPDGVVERARELVEKSDG 397
 DB 228 PVSQDIAEGKWTLMVVKALAESEDERLISILGS-----GDEGSVAEAEIFERYGA 281
 QY 398 LEKTRALAEYAQKALDAIRTFPSPARKALEQUTDKVLTR 438
 DB 282 TOVAHEVALDVVRMAKERLEILEDSDARDALMRITADPVLER 322

RESULT 13
 GGPP_MYCTU STANDARD; PRT; 359 AA.
 AC Q50737;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable geranylgeranyl pyrophosphate synthetase (GGPP synthetase)
 DE (GGPPSAB) (Geranylgeranyl diphosphate synthase) [Includes:
 DE dimethylallyltransferase (EC 2.5.1.1); Geranyltransferase
 DE (EC 2.5.1.10); Farnesyltransferase (EC 2.5.1.29)].
 GN RV3398C OR MT3506 OR MTC78.30.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E., III, Tekaiia F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF
 CC IPP ONTO DMAPP TO FORM GERANYLGERANYL PYROPHOSPHATE WHICH IS A
 CC PRECURSOR OF THE ETHER-LINKED LIPIDS.
 CC -I- CATALYTIC ACTIVITY: dimethylallyl diphosphate + isopentenyl
 CC diphosphate = diphosphate + geranyl diphosphate.
 CC -I- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
 CC = diphosphate + trans-trans-farnesyl diphosphate.
 CC -I- CATALYTIC ACTIVITY: Trans-trans-farnesyl diphosphate + isopentenyl
 CC diphosphate = diphosphate + geranylgeranyl diphosphate.
 CC -I- PATHWAY: BIOSYNTHESIS OF MEMBRANE ETHER-LINKED LIPIDS.
 CC -I- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z77165; BAB01025.1; -
 DR EMBL; AE007156; AAK47843.1; -
 DR TIGR; MT3506; -
 DR TubercuList; RV3398C; -
 DR InterPro; IPR000092; Polyprenyl_synth.
 DR Pfam; PF00348; polyprenyl_synth; 1.
 DR PROSITE; PS00444; POLYPRENYL SYNTHET 2; 1.
 DR PROSITE; PS00723; POLYPRENYL SYNTHET 1; 1.
 KW Hypothetical protein; Lipid synthesis; Isoprene biosynthesis;
 KW Transferase; Complete proteome.
 SQ SEQUENCE 359 AA; 38852 MW; F59F229B19432B6E CRC64;

Query Match 14.1%; Score 309; DB 1; Length 359;
 Best Local Similarity 25.2%; Pred. No. 4.2e-14;
 Matches 102; Conservative 57; Mismatches 162; Indels 84; Gaps 9;

QY 58 EPPNPOLINPLNLVGPMSNLTSNIRSLGSGHPSLDTVAKYV-----OSEG 106
 DB 13 QPDSRMTTRTLPLVGLAHELITFTLRQADRLDPHMRPVVSYHLGWSDERGRPVNNCG 72
 QY 107 KHRPLMWLLMAOATEVAPKQCKEVEVPVNEGLAPPEVLNDKPNMNMRSGLPTKD 166
 DB 73 KAIRPALVFVAEAGADP-----H 92

QY 167 GETEGQTSNIIASQRRLAEITTEMIHAASLLHDDVIDASETRRNAPSQNAFGNKMALAG 226
 DB 93 SAIPGAVS-----VELVHNFSLVHDDLMDRDEHRRHPTVWALWGDMALLAG 140

QY 227 DFLGRASVALARLNPEV---IELLATVIANLVEGEFMQLKNTVDDAIEATATQETFDY 283
 DB 141 DAMLSLAHEVLDCDSPHVGAALRAISEATRELIRGO-----AADTAFE-SRTDVALDE 193

QY 284 YLQKTYLKTASLAKSCRASALGGATPEVADAAAYAGRNGLAFQIVDDMLDYTSATD 343
 DB 194 CLKWAGKTAALMAASAEVAGALLAGAPRSREALVAYGRHIGLAFQIVDDLLGIGWGRPEI 253

QY 344 LGKPAGADLQGLGATAPALFAWKHAEGLPMIKKFS-----PGDVERARELVEKS 395
 DB 254 TGPVVSDDLRSRKTLPVTWTVAHGSGAGRRLAANLWVDETSQTSASDDELAVALIECG 313

QY 396 DGLEKTRALAEYAQKALDAIRT--FPSPARKALEQUTDKVLTR 438
 DB 314 GGRWASAEARRHVTQIGDMVARIGIPDRPAE-LQDLAHYIVDR 357

RESULT 14
 IDSA_METJA STANDARD; PRT; 327 AA.
 AC Q58270;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bifunctional short chain isoprenyl diphosphate synthase [Includes:
 DE Farnesyl pyrophosphate synthetase (EC 2.5.1.1) (FPP synthetase)
 DE (dimethylallyltransferase); Geranyltransferase (EC 2.5.1.10)].
 GN IDSA OR MJ0860.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,

RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.,
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii." Science 273:1058-1073(1996).
RL Science 273:1058-1073(1996).
CC -1- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
CC diphosphate = diphosphate + geranyl diphosphate.
CC -1- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC = diphosphate + trans,trans-farnesyl diphosphate.
CC -1- COFACTOR: REQUIRES MAGNESIUM (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PPP/GGPP SYNTHETASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67530; AAB98865.1; -
CC
CC TRGR: M00860; -
CC
CC DR InterPro: IPR000092; Polyprenyl_synth.
CC DR Pfam: PF00348; Polyprenyl_synth_1.
CC DR PROSITE: PS00444; POLYPRENYL_SYNTHET_2; 1.
CC DR PROSITE: PS00723; POLYPRENYL_SYNTHET_1; 1.
CC KW Lipid synthesis; Isoprene biosynthesis; Transferase; Magnesium;
CC KW Multifunctional enzyme; Complete proteome.
SQ SEQUENCE 327 AA; 37568 MW; 07D68AC9BD57DAC CRC64;

Query Match 13.6%; Score 299; DB 1; Length 327;
Best Local Similarity 29.5%; Pred. No. 1.8e-13;
Matches 78; Conservative 55; Mismatches 111; Indels 20; Gaps 5;

QY 184 AETTEMHAASLHDDVIDASSTRRNAPSGNQAAGKMLAGDPLGRASVALARLN- 242
DB 73 AAVALHNYHTLHDDIMDDDERGRKPTVHVYVGEPMALAGDLVAKAFEVASRIKDN 132
QY 243 ---PEVELLATVIANIIVEGEPQOLKNTVDATATQETFPDYLOKTYLKTASLIAS 239
DB 133 KKAHEVLIKISKACVEVCEGADM-----EFENYTYTMEYTLDMIKTKGALIEAS 184
QY 300 CRASALLGATPEVADAAVYAGNGLAFLQIVDDMLDYVSATD LGKPAQADQLGLATA 359
DB 185 VGAGVADWADNDEEREALEKAYAKIGLTLQIDVDLDLDQKKLGKPGVSDIREGKXTI 244
QY 360 PALFAWKHAH-----LGMITKKKPSDGVARARELVESDGLKTRALAEYAQKALD 414
DB 245 IVIHALKTLEDDKKRLLEILGNKVDEBEIKAEIILKPS--IEVAKELMKQTEBAKE 302
QY 415 AIRFPESPARKLEOLTDVFLVR 438
DB 303 YLKLF-NKDRKRVLEDLADFLMSR 325

RESULT 15
GGPP_SULO STANDARD; PRT; 332 AA.
ID GGPP_SULO
AC P95999;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Geranylgeranyl pyrophosphate synthetase (GGPP synthetase) (GGPS)
DE [includes: Dimethylallyltransferase (EC 2.5.1.1); Farnesyltransferase
DE Geranyltransferase (EC 2.5.1.10); Farnesyltransferase
DE (EC 2.5.1.29)].
GN GDS OR GDS-1 OR SS00061 OR C05010 OR C05_049.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;

RN [1]
RP SEQUENCE FROM N.A. / DSM 1617 / P2;
RC STRAIN=ATCC 35092 /
RX MEDLINE=97055432; PubMed=8899719;
RA Sengen C.W., Klenk H.-P., Singh R.K., Allard G., Chan C.C.-Y.,
RA Liu Q.Y., Penny S.L., Young F., Schenk M.E., Gaasterland T.,
RA Doolittle W.F., Ragan M.A., Charlebois R.L.;
RT "Organizational characteristics and information content of an archaeal
RT genome: 156 kb of sequence from Sulfolobus solfataricus P2.";
RL Mol. Microbiol. 22:175-191(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Awayez M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erasao G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi Ngoc H.P., Redder P., Schenk M.E., Thieriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sengen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -1- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF
CC IPP ONTO DMAPP TO FORM GERANYLGERANYL PYROPHOSPHATE WHICH IS A
CC PRECURSOR OF THE ETHER-LINKED LIPIDS.
CC -1- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
CC diphosphate = diphosphate + geranyl diphosphate.
CC -1- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC = diphosphate + trans,trans-farnesyl diphosphate.
CC -1- CATALYTIC ACTIVITY: Trans-trans-farnesyl diphosphate + isopentenyl
CC diphosphate = diphosphate + geranylgeranyl diphosphate.
CC -1- PATHWAY: BIOSYNTHESIS OF MEMBRANE ETHER-LINKED LIPIDS.
CC -1- SIMILARITY: BELONGS TO THE PPP/GGPP SYNTHETASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y08257; CAAG9541.1; -
CC DR EMBL: AE006647; AAK40423.1; -
CC DR InterPro: IPR000092; Polyprenyl_synth.
CC DR Pfam: PF00348; Polyprenyl_synth_1.
CC DR PROSITE: PS00723; POLYPRENYL_SYNTHET_2; 1.
CC DR PROSITE: PS00444; POLYPRENYL_SYNTHET_1; 1.
CC KW Lipid synthesis; Isoprene biosynthesis; Transferase;
CC KW Multifunctional enzyme; Complete proteome.
CC FT CONFLICT 122 122 A -> S (IN REF. 1).
CC FT CONFLICT 126 126 A -> S (IN REF. 1).
SQ SEQUENCE 332 AA; 37227 MW; 2E3BEC59432CBDA1 CRC64;

Query Match 13.3%; Score 292; DB 1; Length 332;
Best Local Similarity 24.4%; Pred. No. 5.5e-13;
Matches 94; Conservative 67; Mismatches 138; Indels 86; Gaps 9;

QY 76 MSNLTNSIRSLGSHPLDVTAKYVQSEGRKIRPLMVLVLMAGATVAPVQGMKEYE 135
DB 14 VNVNFRKIKNFVSKNFKLEASFFLTFAGKRLRPLVL----- 53
QY 136 VEVNGLAPPEVLNDKNPDMNMRSGPLTKGELIEGQSNILASOR-----RLAETEMIH 191
DB 54 -----SSDLIGDREARAYKAAAVELIH 77
QY 192 AASLLHDDVIDASSTRRNAPSGNQAAGKMLAGDPLGRASVALARLN-----E 244
DB 78 NFTLVHDDIMDDGRRLGLPTVHVYVGEPMALAGDLVAKAFALNALGLDGNFTYK 137
QY 245 VIELLATVIANIIVEGEPQOL--KNTVDATATATQETFPDYLOKTYLKTASLIASGRA 302

Db 138 AFSIFITSIEIISGQMDMSFENRLD-----VTEEE---YIQMIKGTAMLFSCSAAL 188
Qy 303 SALLGATPEVADAAYAYGRNLGLAFOIVDDMLDYTVSATDLGKPGADLQLGLATAPAL 362
Db 189 GGIINKANDDVKKLTEYGLNLGISFQIVDDILGIGDEKELGKPIYSDIREGKKTILVI 248
Qy 363 FAWKHAE-----LGPWKRFSDPGDVERARELVEKSDGLEKTRALAEYAKALDAIR 417
Db 249 KTLSEATEDEKKILVSTLGNKEAKKEDLERASEIIRK-HSLQYAYDLAKKYSDLAENLR 307
Qy 418 TFPES--PARKALEQLTDKVLTRSR 440
Db 308 EIPVSNKTAEKALKYLAQFTIQRK 332

Search completed: January 16, 2003, 09:49:01
Job time : 24 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 08:53:11 ; Search time 67 Seconds

(without alignments)
1353.145 Million cell updates/sec

Title: US-09-830-111A-2

Perfect score: 2194

Sequence: 1 MSPALRISRSRSIASLR.....ESPAKALQLDKVLTRSR 440

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mmc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	823.5	37.5	356	3 P78914	P78914 schizosacch
2	742	33.8	430	5 Q9V923	Q9V923 drosophila
3	734	33.5	416	10 Q9FET8	Q9FET8 quercus rob
4	710.5	32.4	422	10 Q9F789	Q9F789 arabidopsis
5	706	32.2	423	10 Q8RWM1	Q8RWM1 arabidopsis
6	690	31.4	336	11 Q9WU69	Q9WU69 mus musculus
7	683	31.1	376	4 Q9Y2W5	Q9Y2W5 homo sapien
8	678	30.9	321	10 Q9FSW8	Q9FSW8 citrus sine
9	661.5	30.2	323	16 Q8Z0J7	Q8Z0J7 anabaena sp
10	641.5	29.2	393	5 P91093	P91093 caenorhabdi
11	625.5	28.5	245	5 Q8Y08	Q8Y08 drosophila
12	585	26.7	406	10 Q8S948	Q8S948 arabidopsis
13	583.5	26.6	390	10 Q9SYN0	Q9SYN0 arabidopsis
14	574.5	26.2	379	10 Q9SHG4	Q9SHG4 arabidopsis
15	561	25.6	297	10 Q64684	Q64684 arabidopsis
16	550	25.1	363	5 Q964Q8	Q964Q8 trypanosoma

17	455.5	20.8	322	16 Q8XVK7	Q8XVK7 ralistonia s
18	451.5	20.6	323	16 Q8XFR7	Q8XFR7 salmonella
19	451.5	20.6	323	16 Q8X9K0	Q8X9K0 escherichia
20	435	19.8	338	16 Q92RMO	Q92RMO rhizobium m
21	434	19.8	322	2 Q9RBZ8	Q9RBZ8 pseudomonas
22	424.5	19.3	348	16 Q9KUT1	Q9KUT1 vibrio chol
23	420.5	19.2	362	16 Q8YFI8	Q8YFI8 bruceella me
24	418.5	19.1	323	16 Q8ZBAS	Q8ZBAS yersinia pe
25	418.5	19.1	358	16 Q8UHP6	Q8UHP6 agrobacteri
26	416	19.0	336	16 Q9F2X8	Q9F2X8 streptomyce
27	415	18.9	338	16 Q9AG11	Q9AG11 caulobacter
28	411.5	18.8	320	16 Q9RVU0	Q9RVU0 deinococcus
29	407.5	18.6	448	5 Q9VP87	Q9VP87 drosophila
30	405	18.5	322	16 Q9HVL5	Q9HVL5 pseudomonas
31	404.5	18.4	325	2 Q24743	Q24743 rhodobacter
32	402.5	18.3	330	16 Q9CNS9	Q9CNS9 pasteurella
33	402.5	18.3	335	16 Q06428	Q06428 mycobacteri
34	395	18.0	315	2 Q82832	Q82832 gluconobact
35	395	18.0	332	16 Q9PDI8	Q9PDI8 xylella fas
36	381.5	17.4	399	4 Q9NR58	Q9NR58 homo sapien
37	378	17.2	324	16 Q9UQV2	Q9UQV2 neisseria m
38	377.5	17.2	338	16 Q986C2	Q986C2 rhizobium l
39	377	17.2	321	16 Q92A78	Q92A78 listeria in
40	368	16.8	319	16 Q9CFX2	Q9CFX2 lactococcus
41	368	16.8	330	16 Q9CBA5	Q9CBA5 mycobacteri
42	368	16.8	538	5 Q96130	Q96130 plasmodium
43	365.5	16.7	340	17 Q8U1V3	Q8U1V3 pyrococcus
44	363.5	16.6	322	16 Q8R7D2	Q8R7D2 thermoanaer
45	362.5	16.5	319	16 Q99U20	Q99U20 staphylococ

ALIGNMENTS

RESULT 1

P78914 PRELIMINARY; PRT; 356 AA.
AC P78914;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DSC-2001 (TrEMBLrel. 19, Last annotation update)
DE mRNA, partial cds, clone: SV 1754.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PR745;
RX MEDLINE=98162722; PubMed=9501991;
RA Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;
RT "Identification of open reading frames in Schizosaccharomycetes pombe
cDNAs.";
RL DNA Res. 4:363-369(1997).
DR ENBL; D89265; BAA13926.1; -
DR InterPro; IPR000092; Polypreryl synt.
DR Pfam; PF00348; polypreryl synt. 1.
DR PROSITE; PS00723; POLYPRERYL_SYNTHET_1; 1.
DR PROSITE; PS00444; POLYPRERYL_SYNTHET_2; 1.
SQ SEQUENCE 356 AA; 39692 MW; 833865A1780715AE CRC64;

Query Match 37.5%; Score 823.5; DB 3; Length 356;

Best Local Similarity 50.0%; Pred. No. 7.6e-49;

Matches 175; Conservative 56; Mismatches 92; Indels 27; Gaps 4;

QY 67 NPLNLVGPMSNLTSNIRSLGSHPSLDTVAKYVQSEGHIRPLMLVLLMAQATEVAPK 126

Db 32 NASHLIKNELEQISPGIRQLMNSNFEBCSKYTTIAQCKMRPSLVLLMSKATSL--- 88

QY 127 VQWEKVVFPVNEGLAPPEVLNDKNDMMNMRSGPLTKDGEIEGQTSNILLASORLAEI 186

Db 89 FHGIDR-----SVVGDKYTDDDLRS-----FSTGQILPSQLRLAQI 125

```

QY 187 TEMTHAASLHDVYIDASETRRNAPSGNAPGKMAILLADPILGRASVALARLNPVY 246
DB 126 TEMTHIASLHADVIDIAANRRSPSSNVAFGNRRSILANFLIARSTYMAALRNPOVY 185
QY 247 ELATATYANLVEGEFMQKNTVDATAEATATQETFPDYVILQKTLTKTSLIAKSCRSASL 306
DB 186 ELATATYIADLVGEFELQKNTMCPSS-SLEIKQSNFDYIIEKSELKTRASLISKCKRSTIL 244
QY 307 GCATPEVADAAYAVYGRNIGLAFQIVDDMLDYVSATDGLKPGAGADLQGLATAPALPAMK 366
DB 245 GQCSPTATAGAGRGRTIGTAFQIMDDVLDTYSKDTLIGKAAGADLKLGLATAPVLPAMK 304
QY 367 HHAELGPMIKRKSSDPGDVERARELVKSGLEKTRALAEYKOKALDAI 416
DB 305 KYPELGAMIVRFNHPSDIQARSLVECSDAIRPITWALEYIKRPPSV 354

RESULT 2
QY 09V923 PRELIMINARY; PRT; 430 AA.
AC 09V923;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CG3684 protein.
GN CG3684.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Bailew R.M., Basu P.V., Berman B.P., Bhaktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,
RA Jalali M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Maerli B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mialshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacble J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasmaman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
```

```

RL Science 287:2185-2195(2000).
DR EMBL; AE003777; AAF57135.1; -.
DR Flybase; FBgn0039834; CG3684.
DR InterPro; IPR000092; Polyprenyl_synth.
DR Pfam; PF00348; Polyprenyl_synth; 1.
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
SQ SEQUENCE 430 AA; 48759 MW; 3CF533CECF03029 CRC64;

Query Match 33.8%; Score 742; DB 5; Length 430;
Best Local Similarity 37.2%; Pred. No. 4.2e-43;
Matches 167; Conservative 85; Mismatches 135; Indels 62; Gaps 6;
```

```

QY 3 SPALIRIS-SRSIASLR-----SYTLRTASAPSLRLRCTSPSSMAAASARLV 57
DB 33 AGYITNHELTESSKSVRLQPSATKESTSYVQSPHKCSQSLQYSSSKANLQHSVH 92
QY 58 EDPNQPL-----INPLNVGPEMSNLTSNIRSLGSG--HPSLDTVAKYVYQSEGHIRP 111
DB 93 TQCPAGPVRFPQIDPYTLLDDDKFYDVRVYLLKSGTSQPELTTIASYFDGQKALRP 152
QY 112 LMVLLMAQATEVAPRVQGEKVEVPVNEGAPPVINDKNPDMNNRSGPLTRDGEIEG 171
DB 153 NVTMLAKAIXY-----HLNN 168
QY 172 QTSNLAQORRLAETEMTHAASLHDVYIDASETRRNAPSGNAPGKMAILLADPILG 231
DB 169 ESHQVHQRQIALFSEVHSHASLHDVYIDQSDFRKRPVNLMMKKTVMAGDYILS 228
QY 223 RASVALARLNPVEVLELATYANLVEGEFMQKNTVDATAEATATQETFPDYVILQKTLK 291
DB 225 IASIMTARLRSDVTYVLSQILTDLVQSEFMQSGRETE-----NERFMYITKTRK 281
QY 292 TSALIAKSCRSALLGATPEVADAAYAVYGRNIGLAFQIVDDMLDYVSATDGLKPGAG 351
DB 282 TSALIANMLKATAVAIAQADDNVAEVAFGYGRNIGLAFQIVDDMLDFVSTBQMKPTAD 341
QY 352 LQGLATAPALPAMKHAELGPMIKRKSSDPGDVERARELVKSGLEKTRALAEYKOK 411
DB 342 LKLGLATAPVLPACRKYELNPMWRARSEPDVERARELVHSHGLEQTRFLAKKQNE 401
QY 412 ALDAIRTFESPAPKALBOLDKYLTRSR 440
DB 402 AIRLAQELTESPYQKGLQVADVLDVNRMK 430

RESULT 3
QY 09PFT8 PRELIMINARY; PRT; 416 AA.
AC 09PFT8;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Geranyl diphosphat synthase.
GN GPPS.
OS Quercus robur (English oak).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eupsideis I; Fagales; Fagaceae; Quercus.
OX NCBI_TaxID=36942;
[1]
RP SEQUENCE FROM N.A.
RC Zimmer W., Oschinski C.;
RT "Identification of a prenyl transferase gene from Quercus robur.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ298245; CAC20852.1; -.
DR InterPro; IPR000092; Polyprenyl_synth.
DR Pfam; PF00348; Polyprenyl_synth; 1.
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
SQ SEQUENCE 416 AA; 46243 MW; 2914B2D071680AFL CRC64;

Query Match 33.5%; Score 734; DB 10; Length 416;
```

Db	125	KOFRSTILLMATALDV-----RVPEALIGESTDIV-----	155
Qy	167	GEIEGOTSNILASQRRLAEITEMIHAASLLHDDVIDASETRRNAPSGNOAFGNKMAILAG	226
Db	156	-----TSELVRQGIAEITEMIHAVSLLHDDVLDADTRRGVGSINVMGNKMSVLG	209
Qy	227	DFLLGRASVALARLRNPEVIELLATVIANLVEGEFMQLKNTVDDAIEATATQETFDYVLQ	286
Db	210	DELLSRACGALAALKNTVEVALLATAVEHLVTGETMEITS-----STEQRYSMDYMQ	262
Qy	287	KYLYKTASLIANSKRASGALLGGATPEVADAAYAGRNGLGAFQIVDDMLDVTYSATDLGK	346
Db	263	KTYKKTASLISNSCKAVAVLTGQTAEVAVLAFYGRNGLGAFQIIDLTOFTGTSASLGK	322
Qy	347	PAGADQLGLATAPALFAWKHAEALGPMIKRKESDPCDVERARELVEKSDGLEKTRALAE	406
Db	323	GSLSLDIRHGVIATPILFAMEEFPQLEVVDOVEKDPNVDIALEYLGSKGIOQARELAM	382
Qy	407	EVAQKALDAIRTFPESP-----ARKALQOLTQKVLTRSR	440
Db	383	EHANLAAAAIGSLPETDNEDVKRSRRALIDLTHRVITRNK	422
RESULT 5			
ID	Q8RWM1	PRELIMINARY; PRT; 423 AA.	
AC	Q8RWM1;		
DT	01-JUN-2002 (TreeBLrel. 21, Created)		
DT	01-JUN-2002 (TreeBLrel. 21, Last sequence update)		
DE	01-JUN-2002 (TreeBLrel. 21, Last annotation update)		
DE	Putative trans-prenyltransferase.		
OS	AT2G34630.		
GN	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	euroside II; Brassicales; Brassicaceae; Arabidopsids.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,		
RA	Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,		
RA	Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,		
RA	Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,		
RA	Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,		
RA	Ecker J., Theologis A., Davis R.W.		
RRL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY093006; AAM13005.1; --		
KW	Transferase.		
SEQ	SEQUENCE 423 AA; 46514 MW; EE565EI743B781DC CRC64;		
Query Match			
Best Local Similarity 32.2%; Score 706; DB 10; Length 423;			
Matches 178; Conservative 82; Mismatches 133; Indels 68; Gaps			
Qy	1	MASPALRIIRSI--SSRSIASLSRVTLR-----TASAPSLRLRCTPTSRPSSSWAAAVS- 51	
Db	10	ISSKFLRNSFYGSSQSLSHRAFIPTDQGHSCSDSPHKLGVCRITYSLKSPVFGFSH 69	
Qy	52	----SASRLVEPPDNPOLNPLNVLGPEMSNLTNTRLSLLGSGHPSLDTVAKYVQS--E 105	
Db	70	QLYHQSSSLVEEB-----LDPFSLVADELSLLSNKLREMYLAEVPKLSAAAEYFFKRGVQ 124	
Qy	106	GKXIRPLMVLMLAQATEVAPKVGWEKVEVPVNEGLAPPEVLNDKNPDMNMNRSGPLTK 165	
Db	125	GKQFRSTILLMATALNV-----RVPEALIGESTDIV----- 156	
Qy	166	DBIEGOTSNILASQRRLAEITEMIHAASLLHDDVIDASETRRNAPSGNOAFGNKMAILA 225	
Db	157	-----TSELVRQGIAEITEMIHAVSLLHDDVLDADTRRGVGSINVMGNKMSVLA 209	
Qy	226	GDFFLLGRASVALARLRNPEVIELLATVIANLVEGEFMQLKNTVDDAIEATATQETFDYVL 285	

```

Db      210  GDFLLSRACGALAAAKNTEVALLATAVEHLVTGSEITIS-----STEGORSMDYWM 2622
Oy      286  OKTYIKTSLIAKSCRASALLGGAIPREYADAAYAGRNRLGLAFQIVDMDIDYTSAILDGL 3455
Db      263  OKTYKTKTASLISNSCKAAVAVLVTGQTAEEVAVALAFEGRNRLGLAFQIVDIDIDPFGTASL 3222
Oy      346  KPAAGDLOGLATAPALFAMWKHMLGEMIRKKSDDPDVERARELVEKSDGEKETAALA 4055
Db      323  KGSLSIDIRRGVTAITLPAHEEFPQLRREVVOVEKDPENNVIALEVLGSKGIORARELA 3822
Oy      406  EYAOKALDAIRTPESP-----ARKKLEQDTDKVLTLSR 440
Db      383  MEHANLAAAIAGSLPETDNEDYKSRRLIDILTRHVTIRNK 423

```

RESULT 6		
09WU69		
ID	09WU69	PRELIMINARY;
AC	09WU69;	PRT; 336 AA.
DT	01-NOV-1999	(TRENBLrel. 12, Created)
DT	01-NOV-1999	(TRENBLrel. 12, last sequence update)
DT	01-JUN-2002	(TRENBLrel. 21, last annotation update)
DE	Trans-prenyltransferase.	
CN	TPPF OR TPT.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Rocig A., Geromel A., Chretien D., Parfait B., Kadhom N., Edey P.	
RA	Leibaudu M., Rustin P., Munnich A., Appelkvist E.-L., Dallner G.,	
RA	Ernstler L.;	
RT	"Widespread coenzyme Q10 deficiency in familial mitochondrial	
RT	encephalomyopathy.";	
RL	Submitted (Jan-1999)	
RL	to the EMBL/GenBank/DBJ databases.	

RP SEQUENCE FROM N.A.
RA Strausberg R;
RL Submitted (Apr-2002) to the EMBL/Genbank/DBD databases
DR EMBL; AF118855; AAD24462.1; -;
DR EMBL; BC026820; AAD26820.1; -;
DR MGD; MGI:1898278; Tptc.
DR InterPro; IPR000092; PolyPrenyl_synth.
DR Pfam; PF00348; PolyPrenyl_synth. 1.
DR PROSITE; PS00073; POLYPRENYL_SYNTHET_1; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW Transferrase
SQ SEQUENCE 336 AA; 37619 MW; ACEFCC20A955DBD58 CRC64;

Query Match	31.4%;	Score 690;	DB 11;	Length 336;
Best Local Similarity	39.8%;	Pred No. 1.1e-39;		
Matches 149; Conservative	66;	Mismatches 107;	Indels 52;	Gaps 2

67 NPLNLVGPENSLTNI RSL LSGHPSLDTVAKYYVQSEGKHIRPLMLVLLMAQATEVAPK 126

Db 15 DPFKLGWRLKGLYEDIRKELHISTRELKDMSEYYFDGKGKAFRIIVLMARACNIH-- 72

127 VQGWKVEVPVNEG LAPPV LNDKNPDMNMKRSGLTKDGBLEGQTSN LLA SQRLAEI 186

DB / 3 ----- FINNMARK ENVIROCONSULT 09

Country	Year	Value
Algeria	1980	1.00
Algeria	1981	1.00
Algeria	1982	1.00
Algeria	1983	1.00
Algeria	1984	1.00
Algeria	1985	1.00
Algeria	1986	1.00
Algeria	1987	1.00
Algeria	1988	1.00
Algeria	1989	1.00
Algeria	1990	1.00
Algeria	1991	1.00
Algeria	1992	1.00
Algeria	1993	1.00
Algeria	1994	1.00
Algeria	1995	1.00
Algeria	1996	1.00
Algeria	1997	1.00
Algeria	1998	1.00
Algeria	1999	1.00
Algeria	2000	1.00
Algeria	2001	1.00
Algeria	2002	1.00
Algeria	2003	1.00
Algeria	2004	1.00
Algeria	2005	1.00
Algeria	2006	1.00
Algeria	2007	1.00
Algeria	2008	1.00
Algeria	2009	1.00
Algeria	2010	1.00
Algeria	2011	1.00
Algeria	2012	1.00
Algeria	2013	1.00
Algeria	2014	1.00
Algeria	2015	1.00
Algeria	2016	1.00
Algeria	2017	1.00
Algeria	2018	1.00
Algeria	2019	1.00
Algeria	2020	1.00
Algeria	2021	1.00
Algeria	2022	1.00
Algeria	2023	1.00
Algeria	2024	1.00
Algeria	2025	1.00
Algeria	2026	1.00
Algeria	2027	1.00
Algeria	2028	1.00
Algeria	2029	1.00
Algeria	2030	1.00
Algeria	2031	1.00
Algeria	2032	1.00
Algeria	2033	1.00
Algeria	2034	1.00
Algeria	2035	1.00
Algeria	2036	1.00
Algeria	2037	1.00
Algeria	2038	1.00
Algeria	2039	1.00
Algeria	2040	1.00
Algeria	2041	1.00
Algeria	2042	1.00
Algeria	2043	1.00
Algeria	2044	1.00
Algeria	2045	1.00
Algeria	2046	1.00
Algeria	2047	1.00
Algeria	2048	1.00
Algeria	2049	1.00
Algeria	2050	1.00
Algeria	2051	1.00
Algeria	2052	1.00
Algeria	2053	1.00
Algeria	2054	1.00
Algeria	2055	1.00
Algeria	2056	1.00
Algeria	2057	1.00
Algeria	2058	1.00
Algeria	2059	1.00
Algeria	2060	1.00
Algeria	2061	1.00
Algeria	2062	1.00
Algeria	2063	1.00
Algeria	2064	1.00
Algeria	2065	1.00
Algeria	2066	1.00
Algeria	2067	1.00
Algeria	2068	1.00
Algeria	2069	1.00
Algeria	2070	1.00
Algeria	2071	1.00
Algeria	2072	1.00
Algeria	2073	1.00
Algeria	2074	1.00
Algeria	2075	1.00
Algeria	2076	1.00
Algeria	2077	1.00
Algeria	2078	1.00
Algeria	2079	1.00
Algeria	2080	1.00
Algeria	2081	1.00
Algeria	2082	1.00
Algeria	2083	1.00
Algeria	2084	1.00
Algeria	2085	1.00
Algeria	2086	1.00
Algeria	2087	1.00
Algeria	2088	1.00
Algeria	2089	1.00
Algeria	2090	1.00
Algeria	2091	1.00
Algeria	2092	

THE UNITED STATES OF AMERICA
DOCTOR OF THE UNIVERSITY OF CALIFORNIA
AND DAVID E. BROWN, M.D., F.A.C.P., F.A.C.C.
DOCTOR OF THE UNIVERSITY OF CALIFORNIA

307 ССАТБЕУДНАВАУСБНИ СІ ІЕОІУДМІ ДУТІСАТДІ СРБАГОДІ.ОІ СІ ІТАРАІ.ЕАУР 366

203 GCPDPVHETAYOGKNVGTAEOLIDPVLDTSCSDOMGKPTADLKIGTGPVLEACO 262

[illegible]

RESULT 7	
Q9Y2W5	
ID Q9Y2W5	PRELIMINARY;
Source	PRT; 376 AA.

DT 01-NOV-1999 (TREMBL) 12, Created)
DT 01-NOV-1999 (TREMBL) 12, last sequence update)
DT 01-JUN-2001 (TREMBL) 17, last annotation update)
DE trans-prenyltransferase.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxId=9606;

RP SEQUENCE FROM N.A.
RA Roig A., Geomeli V., Chretien D., Parfait B., Kadhon N., Edery P.,
RA Leideard M., Rustin P., Munnich A., Appelkvist E.-L., Dallner G.,
RA Enns L.,
RT "widespread coenzyme Q10 deficiency in familial mitochondrial
RT encephalomyopathy.";
RU Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.

DR InterPro; IPRO00092; Polyiprenyl synt.
DR Pfam; PF00348; Polyiprenyl synt. 1.
DR PROSITE; PS00723; POLYIPRENYL_SYNTHET 1; 1.
DR PROSITE; PS00444; POLYIPRENYL_SYNTHET 2; 1.
KW Transferase.

SQ SEQUENCE 376 AA; 42080 MW; 91B8BE941812D193 CRC64;

Query Match	31.1%	Score	683	DB	4	Length	376
Best Local Similarity	39.8%	Pred. No.	4e-39				
Matches	149	Conservative	65	Mismatches	108	Indels	52
						Gaps	2

67 NP LNLVGPEMSNLITSRSLGSGHPSLDIVAKYVQSEGHIRPLMLVLMQAITEVAPK 126

DD 33 DFFNUGKDUUOLIEGINAEFLISIEUNEISEIIFDVNGNAFNFIIVLIRKACNIN-- 112

[illegible]

EDUCATION OF THE DEAF: A GUIDE TO THE LITERATURE

[illegible]

247 E.I.IATVIANI.VEGEEMOI.KNVDDAIEATATOETEDVYI.OKTYI.KTASI.TAKSCRASAI.I. 306

Db 190 SILTOVIEDLVRGEFLOLGSKE-----NERFAHYLEKTEKKTASTJANSCKAVSVL 242

QY 307 GGATPEVADAAAYAGRNLGLAFQIVDDMLDYTSATDLGKPGADLQGLATAPALFAWK 366

Db 243 GCPDPVHEIAYQYGKNVGIAFQLIDVDLFTSCSDQMGKPTSADLKLGLATGVLFAcq 302

QY 367 HHAELGPMIKRFSDPGDERARELVEKSDGLEKTRALAEYAQKALDAIRTFPESPARK 426

Db 303 QFPEMNAMIMRFS L PGDVDRARQYVLQSDGVQQTYYLAQQYCHEAIRREISKLRPSPERD 362

QY	427	ALBQLTBKVLTRSR	440
		::	:

UD 383 AULGESEI VLIKDA 3 / 6

RESULT 8

ID Q9FSW8 PRELIMINARY; PRT; 321 AA.
 AC Q9FSW8;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Geranyl diphosphate synthase.
 GN GPPS.
 OS Citrus sinensis (Sweet orange).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Sapindales; Rutaceae; Citrus.
 OX NCBI_TaxID=2711;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20521526; PubMed=11069698;
 RA Bouvier F., Suire C., D'Hallange A., Backhaus R., Camara B.;
 RT "Molecular cloning of geranyl diphosphate synthase and
 RL compartmentation of monoterpene synthesis in plant cells.";
 RL Plant J. 24:241-252(2000).
 DR EMBL; AJ243739; CAC16851.1; --
 DR InterPro; IPR000092; Polyprenyl synt.
 DR Pfam; PF00348; polyprenyl synt; 1.
 DR PROSITE; PS00723; POLYPRENYL SYNTHET 1; 1.
 DR PROSITE; PS00444; POLYPRENYL SYNTHET 2; 1.
 DR PROSITE; PS00444; POLYPRENYL SYNTHET 2; 1.
 KW Complete proteome.
 SQ SEQUENCE 321 AA; 35042 MW; 410347CF73C51DBD CRC64;
 [1]
 Query Match 30.9%; Score 678; DB 10; Length 321;
 Best Local Similarity 42.6%; Pred. No. 7e-39;
 Matches 152; Conservative 65; Mismatches 90; Indels 50; Gaps 5;
 QY 92 PSLDTVAKYVQVS--EKGHRLPLMLVLAQAQTEVAPKVGQKVEVPVNEGLAPPEVLN 149
 DB 7 PKLASAEYFFKMGVEGKFRPTVLLMATA-----LN 39
 QY 150 DKQPDMMNMSGPLTKDGEIGQTSNLSASQRLAEITEMIHAASLLHDDVIDASERRN 209
 DB 40 VRPEPH-----DQVEDASTELRTQCCIAEITEMIHWASLLHDDVIDADTRG 91
 QY 210 APGNQAFGNKMAILAGDFLLGRASVALARLNPVEITELATVIANIVEGEFMQKNTVD 269
 DB 92 IGSILNFVGNKGLAVLAGDFLLSRACVALASKNTEVVTLATVVEHLVTGTMQMTSSD 151
 QY 270 DAIEATATQETFDYLYOKTYLKTASLAKSCRASALGGATPEVADAAYAGNGLGAFQ 329
 DB 152 QRC-----SMDYMQKTYKTKASLSNSCKATALLAGQTAEVAILAFDYGNLGLAYQ 204
 QY 330 IVDDMLDYTSATDLGKPGADLQGLATAPALPAWKHAEELGPMIKKPSDPCDVERAR 389
 DB 205 LIDVDLFTGTSASLGKSLSDIRHGIITAPILPAMEEFPQLRVTVVQGFEDSSNDIAL 264
 QY 390 ELVEKSPGLEKTRALAEYAKALDAIRTPPE-----SPARKALEQLTKVLTISR 440
 DB 265 EYLKSRGIQKTRALAVKHANLAAADISLPNNDEVDTKSRRALLDLTHRVITRNK 321
 RESULT 9
 ID Q8Z0J7 PRELIMINARY; PRT; 323 AA.
 AC Q8Z0J7;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Solanesyl diphosphate synthase.
 GN SDS OR ALR0096.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,

RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 DR EMBL; AF003581; BAB77620.1; --
 DR InterPro; IPR000092; Polyprenyl synt.
 DR Pfam; PF00348; polyprenyl synt; 1.
 DR PROSITE; PS00723; POLYPRENYL SYNTHET 1; 1.
 DR PROSITE; PS00444; POLYPRENYL SYNTHET 2; 1.
 KW Complete proteome.
 SQ SEQUENCE 323 AA; 35487 MW; BBE3BB0E1CA91094 CRC64;
 [1]
 Query Match 30.2%; Score 661.5; DB 16; Length 323;
 Best Local Similarity 38.4%; Pred. No. 9.7e-38;
 Matches 145; Conservative 72; Mismatches 102; Indels 59; Gaps 4;
 QY 61 PNQPLINPLNVGPEMSNLTSNIRSLGSGHPSLDTVAKYVQVSEKHKIRPLMLVLAQA 120
 DB 3 PATSLFTP---VEADRLADNLKQLVGNRHPILFAAAEHLFGAGGKRIRPAIVLISRA 59
 QY 121 TEVAPKVGQKVEVPVNEGLAPPEVLNKNPDMNMSGPLTKDGEIGQTSNLSAQ 180
 DB 60 T-----MLDQGITP-----RH 70
 QY 181 RLAEITEMIHAASLLHDDVIDASERTRNAPSGNQAFGNKMAILAGDFLLGRASVALARL 240
 DB 71 RLAEITEMIHTASLVHDDVDSEVRRGVPTVHSLFNGRIAILAGDFLPAQSSWLIANL 130
 QY 241 RNPEVIELLATVIANIVEGEFMQKNTVDIATQETFDYLYOKTYLKTASLAKSC 300
 DB 131 DNQVVKLLSEVIMDIATGEIQGLNRFASI-----SIETIIEKSYKTSASLVANSS 183
 QY 301 RASALGGATPEVADAAYAGNGLGAFQIVDDMLDYTSATDLGKPGADLQGLATAP 360
 DB 184 KAAGLLSEVSPETAELHAYAGRHGLGAFQIVDDILOFTSTDTLGRVSGDLKSGNLTP 243
 QY 361 ALPAWKHAEELGPMIKKPSDPCDVERARELVEKSGGLEKTRALAEYAKALDAIRTPP 420
 DB 244 VLPALAEKPYLEVIREFPAQEGDLEALELIQDSQIQOSRELAHAHTKLAIEHLATLP 303
 QY 421 ESPARKALEQLTKVLTIR 438
 DB 304 PSEHQALIKIAEYAIRS 321
 RESULT 10
 ID P91093 PRELIMINARY; PRT; 393 AA.
 AC P91093;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 43.0 kDa protein.
 GN C24A11.9.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology." the C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Maggi L.;
 RL "The sequence of C. elegans cosmid C24A11.";
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

```

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80443; AAB37678.2; -
DR InterPro; IPR000092; Polyphenyl_synth.
DR Pfam; PF00348; Polyphenyl_synth; 2.
DR PROSITE; PS00723; POLYPHENYL_SYNTHET_1; 2.
DR PROSITE; PS00444; POLYPHENYL_SYNTHET_2; 2.
KW Hypothetical protein.
SQ
SEQUENCE 393 AA; 43003 MW; 4C5FC2EAC87B5E2 CRC64;

Query Match 29.2%; Score 641.5; DB 5; Length 393;
Best Local Similarity 36.7%; Pred. No. 3.1e-36;
Matches 153; Conservative 74; Mismatches 125; Indels 65; Gaps 9;

QY 39 TSSRSSMAAASAS-----RVEPPDPNPPLNPLVGPENSLTSNI 83
DB 22 TSTSTSSSDNSVASTAFVQEHVRQNDIMVQLIPQDESGAVENLADL-----NVTSN- 74
QY 84 RSLIGSGHPSLDTVAKYVVOSEGHIRPLMVLMAQATEVAPKVQGEKVEVVEGLE 143
DB 75 -----LGRNTHYVFOQGGKMLRPVSLMGNA CNSAARSISEVLIAMISTE--- 121
QY 144 PPEVLNDKNPDMNMRSGPLTKDGEIGQTSNIIASGRRLAETEMTHAASLLHDDVIDA 203
DB 122 -----RSG-----IAHLS-VCONQYKIGMAEMHTASLVHDDVIDE 158
QY 204 SETRRNAPSGNQAFGNKMAIIAGDPLIGRASVALARLNEVEVELLATVIANVEGEFMQ 253
DB 159 ANRRRGASVVAAGNKMVSVLVGFILARATQILCSIGKPIILSVKMSIIEDLVGEFMQ 218
QY 264 LKATVDATETATQETFDYVLOKTYLKTASLIKSCRASALLG-GATPEVADAAYVGR 322
DB 219 MSTTPTATPV---DKMKAYIEKTHKRTASLPASSCRSAIILADSDLLHETAPFYGR 274
QY 323 NLGIAPQIVDMLDYTVASATDLGKPAQADLQLGATAPALFAMKHAELGPMIKRFSDP 382
DB 275 NLGIAPQIADLDLDFITADEMGKPVADLKLGLATAPVLVACEQYELMTMLRKPKHD 334
QY 383 GDVERARELVKSDGLEKTRALAEVYQKALDAIRTPESPARKALEQLDIXVLTS 439
DB 335 GDAEKAREIVNSDGMKTRLLIDSYQKAVEMASSLPN--RNSETHLIKILMASOS 389

RESULT 11
Q88Y08 PRELIMINARY; PRT; 245 AA.
AC Q88Y08;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, last annotation update)
DE REL8374P.
GN CG3684.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Chame M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nuno J., Pacieb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celinker S.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY075463; AAL68276.1; -
SQ
SEQUENCE 245 AA; 27384 MW; 66CCD7DD5A8A3A1D CRC64;

Query Match 28.5%; Score 625.5; DB 5; Length 245;
Best Local Similarity 49.6%; Pred. No. 1.9e-35;
Matches 125; Conservative 51; Mismatches 69; Indels 7; Gaps 1;

QY 189 MIHAASLLHDDVIDASERRRNAPSGNQAFGNKMAIIAGDPLIGRASVALARLNEVELEL 248
DB 1 MHSASLVHDDVIDIYDSDRRGKPSNALMNHKTKTMAGDYILSTASTIMANLRSDDTIV 60
QY 249 LATVIAMIVEGEFMQLKNTVDAIEATQETFDYVLOKTYLKTASLIKSCRASALLG 308
DB 61 LQIITLDVQGEFMQLGSRETE-----NEPFAVYLTKYTKTASLIANLAKTAVIAQ 113
QY 309 ATPEVADAAYAGRNLGLAFQIVDMLDYTVASATDLGKPAQADLQLGATAPALFAMKHN 368
DB 114 ADDNVAEVAFOYGRNIGLAFQIVDMLDVFSSTQMGKPTAADLKLGLATAPVLFACEKY 173
QY 369 AELGPMIKRFSDPGDVRRARELVKSDGLEKTRALAEVYQKALDAIRTPESPARKAL 428
DB 174 PELNPMWRFRFEPDVRARELVKSHGLEQTRFLAKKCNEMAIRLAQELTESPYQXGL 233
QY 429 EQLTDKVLTRSR 440
DB 234 QVADLVINRMK 245

RESULT 12
Q88948 PRELIMINARY; PRT; 406 AA.
AC Q88948;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, last annotation update)
DE Polyphenyl diphosphate synthase.
GN AT-TRANS-PT.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLOMBIA;
RA Hirooka K., Fukusaki E., Kobayashi A.;
RT "Polyphenyl Diphosphate Synthase Involved in Quinine Biosynthesis from
RT Higher Plant.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB071514; BAB86941.1; -
SQ
SEQUENCE 406 AA; 44468 MW; 2FDD1877A569411 CRC64;

Query Match 26.7%; Score 585; DB 10; Length 406;
Best Local Similarity 33.6%; Pred. No. 2.6e-32;
Matches 135; Conservative 69; Mismatches 140; Indels 58; Gaps 5;

QY 35 RCTPTSRSSSWMAAASASRL---VEPDPNP--LNPPLVGPENSLTSNIRSLIGS 89
DB 49 KAVPTSKSEIISLNGISQTVSPDLKQESKQPSLVTLFELVAVDLQTLNDNLSTVGA 108
QY 90 GHPSLDTVAKYVVOSEGHIRPLMVLMAQATEVAPKVQGEKVEVVEVNEGLAPPEVLN 149
DB 109 ENPVLISAAGQIFGAGGRMRGLVYSHAT----- 140
QY 150 DKNPDMNMRSGPLTKDGEIGQTSNIIASGRRLAETEMTHAASLLHDDVIDASERRN 209
DB 141 -----AELAG-LKELTERRRLAETEMTHAASLLHDDVIDESDMRG 182
QY 210 ASGNQAFGNKMAIIAGDPLIGRASVALARLNEVEVELLATVIANVEGEFMQLKTV 269
DB 183 KETVELEGTGTVAVLAGSPFAQASWYLANLENLEVKLIQVLIKDFASGEIKQASSLFD 242
QY 270 DATETATQETFDYVLOKTYLKTASLIKSCRASALLGATPEVADAAYAGRNLGLAFQ 329
DB 243 -----CDTKLDEYLLSKFYKTASLVASTKGAATFRRVPEVDTQMYERGKXGLSLFQ 295
```

```
QY 330 IVDDMLDYTVSATDLGKPGADLQGLATAPALFAWKHAEIAGPMIKRKFSDFGDEVERAR 389
DB 296 IVDDILDTOSTQOLGKPGAGSLAKGNLTAPVIFALEREPLREIIESECFEAGSLEAEI 355
QY 390 ELVEKSDGLEKTRALAEYAKALDAIRTPESPARKALEOL 431
DB 356 EAVTKGGGIKRAELAREKADDAIKNLQCLPRSGFRSALED 397

RESULT 13
QSYNO PRELIMINARY; PRT; 390 AA.
AC Q9SYNO;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE T30F21.15 protein.
GN T30F21.15.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007260; AAD30584.1; -.
DR InterPro; IPR000092; Polyprenyl_synth.
DR Pfam; PF00348; polyprenyl_synth; 1.
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
SQ SEQUENCE 390 AA; 43488 MW; 8941EC6602DA8155 CRC64;

Query Match 26.6%; Score 583.5; DB 10; Length 390;
Best Local Similarity 35.0%; Pred. No. 3.1e-32;
Matches 132; Conservative 65; Mismatches 125; Indels 55; Gaps 4;

QY 57 VEPDPNPQ--LIPLNIVGPEMSNLTNIRSLGSGHPSLDTVAKYVQSEGKHIRPLMV 114
DB 58 LKQESKQPIISIVTLFELVAVDQLNDNLLSIVGAENPVLISAAEQIFGAGGKMRPLV 117
QY 115 LLMAQATEVAPKVGQEKVEVPVNEGLAPPEVLNDKNPDMNMRSGPLTKDGEIEGQTS 174
DB 118 FLVSHAT-----LDDYMLKSYKYKTASLVAASTKGAIFSKVE 249
QY 175 NILASORLAEITEMIHASILLHDDVIDASETRNPASGNQAFGNKMAILAGDFLLGRAS 234
DB 132 ELTEHRLAEITEMIHASILLHDDVIDESDMRGKTFVHELFGTRVAVLAGDFMFAQAS 191
QY 235 VALARLNPEVIELLATVIANLVEGEFMQKNTVDDAIEATATQETPDYLYLQKTYLKTAS 294
DB 192 WYLANLENLEVIKLSIQVIXDFASGEIKQASSLFD-----CDTKLDEYLLKSYFTAS 244
QY 295 LIAKCRASALLGGATEVADAAYAGRNGLAFQIVDDMLDYTVSATDLGKPGADLQOL 354
DB 245 LVAASTKGAIFSRVDPVTQMEYFEGKNLGLSFQIVDDILDFTQSTQOLGKPGAGSLAK 304
QY 355 GLATAPALFAWKHAEIAGPMIKRKFSDFGDEVERARELVEKSDGLEKTRALAEYAKALD 414
DB 305 GNLTAPVIFALEREPLREIIESECFEAGSLEAEIAVTKGGGIKRAQELAREKADDAIK 364
QY 415 AIRTFPESPARKALEOL 431
DB 365 NUQCLPRSGFRSALED 381
```

```
RESULT 14
Q9SHG4 PRELIMINARY; PRT; 379 AA.
AC Q9SHG4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Very similar to prenyl transferase.
GN F20D23.25.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007651; RAD50025.1; -.
DR InterPro; IPR000092; Polyprenyl_synth.
DR Pfam; PF00348; polyprenyl_synth; 1.
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW Transferase.
SQ SEQUENCE 379 AA; 41916 MW; AC07C095915D989E CRC64;

Query Match 26.2%; Score 574.5; DB 10; Length 379;
Best Local Similarity 35.2%; Pred. No. 1.2e-31;
Matches 127; Conservative 66; Mismatches 115; Indels 53; Gaps 3;

QY 71 LVCPENSLTNSIRSLGSGHPSLDTVAKYVQSEGKHIRPLMVLLMAQATEVAPKVGW 130
DB 63 VWADDLQRLNDNLLSIVGAENPVLISAAEQIFGAGGKMRPLGVFLVSRAT----- 113
QY 131 EKVEVPVNEGLAPPEVLNDKNPDMNMRSGPLTKDGEIEGQTSNILASORRLAEITEMI 190
DB 114 -----AELAG-IKEITVEHRLGEIEMI 136
QY 191 HAASLLHDDVIDASETRNPASGNQAFGNKMAILAGDFLLGRASVALARLNPEVIELLA 250
DB 137 HTASLIHDDVIDESDMRGKTFVHELFGTRVAVLAGDFMFAQASWYLANLENLEVIKLS 196
QY 251 TVTANLVEGEFMQKNTVDDAIEATATQETPDYLYLQKTYLKTASLAKCRASALLGGAT 310
DB 197 QVTKDFASGEIKQASSLFDCKVK-----LDDYMLKSYKYKTASLVAASTKGAIFSKVE 249
QY 311 PEVADAAYAGRNGLAFQIVDDMLDYTVSATDLGKPGADLQGLATAPALFAWKHAE 370
DB 250 SKVAEQMYQFGKNLGLSFQVDDILDFTOSTQOLGKPAANDLAKGNITAPVIFALENEPR 309
QY 371 LGPMIKRKFSDFGDEVERARELVEKSDGLEKTRALAEYAKALDAIRTFPESPARKALEQ 430
DB 310 LREIIESECFEAGSLEAEIIEVNRGGIKKAQELAKEKALKNLNCILPRSGFRSALED 369
QY 431 L 431
DB 370 M 370

RESULT 15
O64684 PRELIMINARY; PRT; 297 AA.
AC O64684;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Putative trans-prenyltransferase.
GN AT2G34630.
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 16, 2003, 06:38:46 ; Search time 2939 Seconds
(without alignments)
16368.472 Million cell updates/sec

Title: US-09-830-111A-1
Perfect score: 1653
Sequence: 1 ttttggggtcgaagtc.....atgcgaactcagaagtga 1653

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : GenEmbl.*
- 1: gb.ba.*
 - 2: gb.btg.*
 - 3: gb.in.*
 - 4: gb.om.*
 - 5: gb.ov.*
 - 6: gb.pat.*
 - 7: gb.ph.*
 - 8: gb.pl.*
 - 9: gb.pr.*
 - 10: gb.ro.*
 - 11: gb.sts.*
 - 12: gb.sy.*
 - 13: gb.un.*
 - 14: gb.vi.*
 - 15: em.ba.*
 - 16: em.fun.*
 - 17: em.in.*
 - 18: em.hum.*
 - 19: em.mu.*
 - 20: em.or.*
 - 21: em.ov.*
 - 22: em.pat.*
 - 23: em.ph.*
 - 24: em.pl.*
 - 25: em.ro.*
 - 26: em.sts.*
 - 27: em.un.*
 - 28: em.vi.*
 - 29: em.htg.hum.*
 - 30: em.htg.inv.*
 - 31: em.htg.other.*
 - 32: em.htg.mus.*
 - 33: em.htg.pln.*
 - 34: em.htg.rod.*
 - 35: em.htg.mam.*
 - 36: em.htg.vit.*
 - 37: em.sy.*
 - 38: em.htgo.hum.*
 - 39: em.htgo.mus.*
 - 40: em.htgo.other.*
 - 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1653	100.0	1653	6	E51323	E51323 Process for
2	242.8	14.7	2214	8	D84311	D84311 Schizosacch
3	231.4	14.0	2666	8	MC1496300	AJ496300 Mucor cir
4	222	13.4	1134	6	E13333	E13333 gDNA encodi
5	218.6	13.2	1492	8	D89265	D89265 Schizosacch
6	199.6	12.1	2097	8	SPBPU694	AL136498 S.pombe c
7	198.6	12.0	5371	12	CVU62637	U62637 Cloning vec
8	198	12.0	4176	6	A13038	A13038 Artificial
9	198	12.0	4176	12	SYNTRC99A	M27744 Cloning vec
10	198	12.0	4176	12	XXU13872	U13872 pTtc99a clo
11	198	12.0	4239	6	A29289	A29289 pSEC-Bpl co
12	197	11.9	4613	12	ASPBAD18	X81838 E.coli DNA
13	197	11.9	5075	12	U02444	U02444 Cloning vec
14	186.4	11.3	2799	3	AY075463	AY075463 Drosophil
15	186.2	11.3	4542	12	ASPBAD24	X81837 E.coli DNA
16	186.2	11.3	6775	12	AV112733	AV112733 Broad hos
17	185.6	11.2	1262	10	AF118855	AF118855 Mus muscu
18	185.6	11.2	1552	10	BC026820	BC026820 Mus muscu
19	185.4	11.2	6124	12	AY112734	AY112734 Broad hos
20	184	11.1	4232	12	SYNLACIQ	L24193 Cloning vec
21	183	11.1	4735	12	SYNBETALAC	L24367 Cloning vec
22	182.8	11.1	1600	8	ORO298245	AJ298245 Quercus r
23	182.4	11.0	1402	8	ATH17376	Y17376 Arabidopsis
24	182.4	11.0	1408	8	AY093006	AY093006 Arabidops
25	180.8	10.9	966	6	AX043804	AX043804 Sequence
26	180.8	10.9	966	6	AX360800	AX360800 Sequence
27	180.8	10.9	966	6	AX428168	AX428168 Sequence
28	180	10.9	4991	12	EVPK2333	X95387 Expression
29	179.8	10.9	1114	8	CSP243739	AJ243739 Citrus si
30	178	10.8	2681	12	SYNPUT18	D13250 Plasmid put
31	171.8	10.4	1485	9	AF118395	AF118395 Homo sapi
32	171.8	10.4	298750	1	AP005375	AP005375 Thermosyn
33	168.6	10.2	1641	9	AK024802	AK024802 Homo sapi
34	165	10.0	5241	6	AX027815	AX027815 Sequence
35	160.4	9.7	3302	6	A20357	A20357 plasmid pIG
36	160.4	9.7	3302	6	A20358	A20358 plasmid pIG
37	159.6	9.7	133859	1	D90899	D90899 Synchocyst
38	155.2	9.4	4009	6	A39734	A39734 Sequence 2
39	155.2	9.4	4009	6	AR069625	AR069625 Sequence
40	152.4	9.2	3423	6	AR9078	AR9078 Sequence 86
41	152.4	9.2	3423	6	ARI95392	ARI95392 Sequence
42	152.4	9.2	4633	6	A20355	A20355 plasmid pIG
43	152.4	9.2	4633	6	A20356	A20356 plasmid pIG
44	152	9.2	1097	6	AR008977	AR008977 Sequence
45	152	9.2	1097	6	AR075318	AR075318 Sequence

ALIGNMENTS

RESULT 1
E51323
LOCUS E51323
DEFINITION Process for producing coenzyme Q10.
ACCESSION E51323
VERSION E51323.1 GI:18633579
KEYWORDS JP 2001061478-A/1.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1653)
AUTHORS Matsuda,H., Kawamukai,M., Vajima,A., Ikenaka,Y., Hasegawa,J. and Takahashi,S.
TITLE Process for producing coenzyme Q10
JOURNAL Patent: JP 2001061478-A 1 13-MAR-2001;

E51323 1653 bp DNA linear PAT 31-JAN-2002

COMMENT KANEKA CORP
OS Saioella complicata
PN JP 2001061478-A/1
PD 13-MAR-2001
PF 24-AUG-1999 JP 1999237561
PR
PI HIDEYUKI MATSUDA, MAKOTO KAMAMUKAI, AKIYOSHI YAJIMA, PI
YASUHIRO IKENAKA,
PI JUNZO HASEGAWA, SATOMI TAKAHASHI
PC C12N15/09, C12N1/21, C12N9/12, C12N1/21, C12R1:19, (C12N9/12, PC
C12R1:19)
PC C12N15/00
CC
FH Key Location/Qualifiers
FT source 1. 1653
FT /organism='Saioella complicata'.
FEATURES
source 1. 1653
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 392 a 417 c 475 g 369 t
ORIGIN

Query Match 100.0%; Score 1653; DB 6; Length 1653;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTGGGGGTGGAAGAATCGGACCGGTCAGAGTTCCGCTTGAGACCGTAAGGCTCG 60
DB 1 TTTTGGGGGTGGAAGAATCGGACCGGTCAGAGTTCCGCTTGAGACCGTAAGGCTCG 60
QY 61 GAGATTGAGTTCAGAGCAAAAGCTTTGATCCGTGAGGTTTACATCTTCAGCAAAATCTTTC 120
DB 61 GAGATTGAGTTCAGAGCAAAAGCTTTGATCCGTGAGGTTTACATCTTCAGCAAAATCTTTC 120
QY 121 AATCCATATACCATGAGCTTCACACGACCTCGGAGATACGAGATAGCTTTCGATCAT 180
DB 121 AATCCATATACCATGAGCTTCACACGACCTCGGAGATACGAGATAGCTTTCGATCAT 180
QY 121 AATCCATATACCATGAGCTTCACACGACCTCGGAGATACGAGATAGCTTTCGATCAT 180
DB 121 AATCCATATACCATGAGCTTCACACGACCTCGGAGATACGAGATAGCTTTCGATCAT 180
QY 181 GCGCTCTCGGATCGGTTTACCTTAAGAACAGCTCGGACCTTCATTAAGCAATAGT 240
DB 181 GCGCTCTCGGATCGGTTTACCTTAAGAACAGCTCGGACCTTCATTAAGCAATAGT 240
QY 241 TACCCCGACGAGCCCGGATTCAGATTCAGTGGCTGCTGTGTCTTCGCGCTGAGACT 300
DB 241 TACCCCGACGAGCCCGGATTCAGATTCAGTGGCTGCTGTGTCTTCGCGCTGAGACT 300
QY 301 GGTGAGCTGATCCGATCAACCTCTCATCAATCCGCTCAACTGGTGGTCCGAGAT 360
DB 301 GGTGAGCTGATCCGATCAACCTCTCATCAATCCGCTCAACTGGTGGTCCGAGAT 360
QY 361 GTCGAATCTTAATCCAAATCCGATCTCTCCGTTGAGACACCTTCCTCGACAC 420
DB 361 GTCGAATCTTAATCCAAATCCGATCTCTCCGTTGAGACACCTTCCTCGACAC 420
QY 421 TGTCCCTAAATATATATGTTCACTGAGGGAAGCATATTCGTCGCTCATGGTACTGCT 480
DB 421 TGTCCCTAAATATATATGTTCACTGAGGGAAGCATATTCGTCGCTCATGGTACTGCT 480
QY 481 GATGCTCAGCGACGAGGATTCGCAAAAGTTTCAAGGTTGGAGAGAGTGGTGGAGGT 540
DB 481 GATGCTCAGCGACGAGGATTCGCAAAAGTTTCAAGGTTGGAGAGAGTGGTGGAGGT 540
QY 541 TCCGCTGAAACGAGGATTCGCAAAAGTTTCAAGGTTGGAGAGAGTGGTGGAGGT 600
DB 541 TCCGCTGAAACGAGGATTCGCAAAAGTTTCAAGGTTGGAGAGAGTGGTGGAGGT 600
QY 601 GAACATGAGTTCAGAGCACTTAAGCAAGAGCGGATTCAGAGGAGAGAGTTCGATAT 660
DB 601 GAACATGAGTTCAGAGCACTTAAGCAAGAGCGGATTCAGAGGAGAGAGTTCGATAT 660
QY 661 CCTGCTTCGCAACGCGGTTGGCTGAGATTCAGAGATGATTCATGACATCACTCTCT 720
DB 661 CCTGCTTCGCAACGCGGTTGGCTGAGATTCAGAGATGATTCATGACATCACTCTCT 720

DB 661 CCTGCTTCGCAACGCGGTTGGCTGAGATTCAGAGATGATTCATGACATCACTCTCT 720
QY 721 CCACGACGACGTTATCGACGCTTCGAGACGAGAAACGACCATTCGGAACCGAGC 780
DB 721 CCACGACGACGTTATCGACGCTTCGAGACGAGAAACGACCATTCGGAACCGAGC 780
QY 781 ATTGGAAACAAAGATGAGGATTTTGGCTGATTTCTTGTGGGACGGGCTGTGTTC 840
DB 781 ATTGGAAACAAAGATGAGGATTTTGGCTGATTTCTTGTGGGACGGGCTGTGTTC 840
QY 841 ATTGGCAAGTTGGCAATCCGAGAGTATGAGCTTTGGCTACTGTATTTGCAAACTT 900
DB 841 ATTGGCAAGTTGGCAATCCGAGAGTATGAGCTTTGGCTACTGTATTTGCAAACTT 900
QY 901 GGTGAGGAGATTCATGACATTCGAAATACTTTGATGATGATGATGATGATGATGATG 960
DB 901 GGTGAGGAGATTCATGACATTCGAAATACTTTGATGATGATGATGATGATGATGATG 960
QY 961 GACGAGGAAACGTTGATTCATTTGAGAGACCTTACTTAAGACTGCTGCTGAT 1020
DB 961 GACGAGGAAACGTTGATTCATTTGAGAGACCTTACTTAAGACTGCTGCTGAT 1020
QY 1021 TGGCAAGTCGACAGAGAGGAGGCTTCGGGTGGTCTAGGCTGAGTGGTGGTGGTGGT 1080
DB 1021 TGGCAAGTCGACAGAGAGGAGGCTTCGGGTGGTCTAGGCTGAGTGGTGGTGGTGGT 1080
QY 1081 TGTCTTATCTTACGGAAGAACCTTTGATTCGATTCGATTCGATTCGATTCGATTCG 1140
DB 1081 TGTCTTATCTTACGGAAGAACCTTTGATTCGATTCGATTCGATTCGATTCGATTCG 1140
QY 1141 CTACACCGCTTCGCTTACGACCTTCGATTCGATTCGATTCGATTCGATTCGATTCG 1200
DB 1141 CTACACCGCTTCGCTTACGACCTTCGATTCGATTCGATTCGATTCGATTCGATTCG 1200
QY 1201 GCGCAGCGGCGGCGCTTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT 1260
DB 1201 GCGCAGCGGCGGCGCTTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT 1260
QY 1261 GCGCAAGTTCTCTGACCCGAGAGAGCTGACGAGCTTCGATTCGATTCGATTCGAT 1320
DB 1261 GCGCAAGTTCTCTGACCCGAGAGAGCTGACGAGCTTCGATTCGATTCGATTCGAT 1320
QY 1321 TGGATTGAGAAAGACGAGAGCTTCGAGAGAGTATCCGAGAGGCTTGGATGAT 1380
DB 1321 TGGATTGAGAAAGACGAGAGCTTCGAGAGAGTATCCGAGAGGCTTGGATGAT 1380
QY 1381 TGGAGCTTCGCGAGAGTCCGAGAGGCTTGGAGAGTTCGAGAGGCTTGGATGAT 1440
DB 1381 TGGAGCTTCGCGAGAGTCCGAGAGGCTTGGAGAGTTCGAGAGGCTTGGATGAT 1440
QY 1441 GACTAGTCAAGATGAGATTCGAGCTCGTACCGGAGATTCCTTGAAGTTCGACTGCA 1500
DB 1441 GACTAGTCAAGATGAGATTCGAGCTCGTACCGGAGATTCCTTGAAGTTCGACTGCA 1500
QY 1501 GGCATGCAAGCTTGGCTTTTGGCGATGAGAGAGATTTTCAGCTCATAGATTA 1560
DB 1501 GGCATGCAAGCTTGGCTTTTGGCGATGAGAGAGATTTTCAGCTCATAGATTA 1560
QY 1561 ATCAAGACGAGAGAGGCTTCGATTAACAGAAATTTGGCTGCGGACATGAGCGGTGT 1620
DB 1561 ATCAAGACGAGAGAGGCTTCGATTAACAGAAATTTGGCTGCGGACATGAGCGGTGT 1620
QY 1621 CCACCTGACCCGATGCGCAACTCAGAGTGA 1653
DB 1621 CCACCTGACCCGATGCGCAACTCAGAGTGA 1653

RESULT 2
D84311 D84311 2214 bp DNA linear PLN 06-FEB-1999
LOCUS Schizosaccharomyces pombe gene for decaprenyl diphosphate synthase,
DEFINITION complete cds.
ACCESSION D84311

D84311.1 GI:1845554
 dps; decaprenyl diphosphate synthase; ubiquinone biosynthesis.
 Schizosaccharomyces pombe DNA.
 Schizosaccharomyces pombe
 Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 Schizosaccharomycetales; Schizosaccharomycetaceae;
 Schizosaccharomycetes.
 1 (sites)
 Suzuki, K., Okada, K., Kamiya, Y., Zhu, X.F., Nakagawa, T., Kawamukai, M.
 and Matsuda, H.
 Analysis of the decaprenyl diphosphate synthase (dps) gene in
 fission yeast suggests a role of ubiquinone as an antioxidant
 J. Biochem. 121 (3), 496-505 (1997)
 97279041
 2 (bases 1 to 2214)
 Direct Submission
 Suzuki, K.
 Submitted (03-APR-1996) Kengo Suzuki, Shimane University, Faculty
 of Life and Environmental Science, Nishikawatsu 1060, Matsue,
 Shimane 690, Japan (E-mail:kawamuka@botan.shimane-u.ac.jp,
 Tel:0852-32-6587, Fax:0852-32-6587)
 On Feb 21, 1997 this sequence version replaced gi:1304267.
 Sequence updated (19-Feb-1997) by:Kengo Suzuki.
 Location/Qualifiers
 1. .2214
 /organism="Schizosaccharomyces pombe"
 /db_xref="taxon:4896"
 347. .1483
 /gene="dps"
 347. .1483
 /gene="dps"
 /codon_start=1
 /product="decaprenyl diphosphate synthase"
 /protein_id="BAA12314.1"
 /db_xref="GI:1845555"
 /translation="MTQYVLYKMRKLSLKVRSVTLRSTNRNASHLKNLEQI
 SGRIOMNSSEFLIECKYTIAGCKOMRPSVLMSKATSLCHGIDRSVVDKRYI
 DDDKRSFTGQLPSQLAQIEMHIALSHLDHVIDHANVRGSPSSNVAFGNRR
 SILAGNFIARASTAMARLNQVTELLATADLVRFGLQKNTMFPDSLEIKQSN
 FVYIEKSLKASTLSKSGASTILGQCSPTVATAGRGICGTAFQMDVLDYI
 SKDDTLKGAAGADLKLATAPVLPAWKYPPELGAMIVNREHNPSPDIQRSLVETD
 ALEQITWAKYIKAKDSILCLPDSFARKALPALADKVIIRKK"
 BASE COUNT 676 a 408 c 403 g 727 t
 ORIGIN
 Query Match 14.7%; Score 242.8; DB 8; Length 2214;
 Best Local Similarity 57.4%; Pred. No. 6.6e-42;
 Matches 458; Conservative 0; Mismatches 337; Indels 3; Gaps 1;
 QY 649 GAGTCGAATATCTCGCCTCGCAACGGCGTTGGCTGAGATCAGGAGATGATCCATGC 708
 DB 679 GACGGGTCAAATCTTCTCTCAATTGAGATTAGCACAATAACCGAGATGATCCATAT 738
 QY 709 AGCATCACTCTCCACACGACGTTATCGAGCTTCGAGACGACGAAACGACCATC 768
 DB 739 AGCAAGTTTCTGTCATGACCATGTGATGATCAACGCTTAATGTCCGTAGAGCTCACCTTC 798
 QY 769 CGGAACACGACGATTCGGAACCAAGATGGCGATTTTCGGCTGGTGATTTCTTTGGGACG 828
 DB 799 AAGCAATGTTGCTTCGGTAAATCGACGGTCAATCTTTCGGGTAATTTTCATCTTCGACG 858
 QY 829 GGCGTCTGTTGCAATGGCGAGGTTGCGCAATCCGGAGGTGATTGAGCTTTTGGCTACTGT 888
 DB 859 GGCTTCGACTGCTATGCGCCGCTTCGAAATCCCAAGTTACGGAGTTGTTAGTCAAGT 918
 QY 889 TATTGCAACTTGGTTAGGGAGGTTTCATGCACTGAAAAAATCTGTTGATGCGAT 948
 DB 919 GATAGCAGACTTGGTTTCGAGGTGAGTTTTCGAGCTAAAAAATACTATGATCCTTCATC 978
 QY 949 TGAGGCTACGCGACGACGAAACGTTTCGATTACTATTTCGAGAGACTTACTTGAAGAC 1008
 DB 979 TTTGGAAA---TAAACAATCAAAATTTTGACTATTATATGAAAAAAGTTTTTGAAGAC 1035

QY 1009 TCGTCTCTTATTCGCAAGTCGTGAGCAAGTCGCTTCTGGTGGTCTAGCCTGA 1068
 DB 1036 AGCGAGTTAAATTTCCAAAAGCTGCAAGGCTTCTACAATCTTCGACAATGTTCTCTAC 1095
 QY 1069 GGTGCTGATGCTGCTTATGCTTACGGAAGAACCTTTGGTTTGGCATTTCCAGATCGTCGA 1128
 DB 1096 TGAGCAACAGCTGCTGGAGAATACGGTCGATGCTTGTACTGCTTTTCAACTAATGGA 1155
 QY 1129 CGACATGCTGACTACACCGTCTCCGCTACCGACCTCGGTAAAGCCCGCGGTGAGACCT 1188
 DB 1156 TGACGTTGTTGGACTATACGTCGAAAGATGATCTTTTAGGAAAGCGGCTGTTGAGATTT 1215
 QY 1189 CAGCTCGGCTCGCCACCGCGGCGCTCTTCGATGGAAGCACCAACCGCGAGCTCGG 1248
 DB 1216 GAAGCTAGGGTTGGCTACAGCTCCCGTCTCTTTGTCATGAAAAGATATCCAGAACTGG 1275
 QY 1249 TCCCATGATCAACGCGAAGTTCTCTGACCCAGGAGAGCTCGAGCGTGCAGCGAGTTGGT 1308
 DB 1276 TGAATGATGTAAGATGATTCATCTCTGATATCAACGGGCTGTTCTTTGGT 1335
 QY 1309 CGAGAAAAGTGAATGATGAGAGAGAGAGAGCTTGGCGGAGAGATATCCCAAGAGGC 1368
 DB 1336 TGAGTGCACTGATGCTATCGAGCAACCATCACTTGGGCAAAAAGATATCAAAAAGC 1395
 QY 1369 GTTGGATCAATTCGACGCTTCCGAGAGCTCCGCGACGAGCTTTGGAGCAGTTGAC 1428
 DB 1396 CAAAGATTCCTTCTGTCTCTCCCTGATTCACCTGCAAGGAGGCACTTTTTCGTTGGC 1455
 QY 1429 GGACAAAGTGTTCAGTAG 1446
 DB 1456 TGATAAGTAATAACGAG 1473
 RESULT 3
 MCI496300
 LOCUS
 DEFINITION
 Mucor circinelloides f. lusitanicus isob gene for solanesyl
 pyrophosphate synthase, exons 1-3.
 ACCESSION
 AJ496300
 VERSION
 AJ496300.1 GI:21955861
 KEYWORDS
 isob gene; solanesyl pyrophosphate synthase.
 Mucor circinelloides f. lusitanicus.
 Mucor circinelloides f. lusitanicus
 Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
 Mucor.
 ORGANISM
 Velavos, A.
 Carotenogenesis in Mucor circinelloides
 Thesis (2001) Department of Microbiologia y Genetica, Universidad
 de Salamanca, Salamanca, Spain
 REFERENCE
 1
 Velavos, A.
 Carotenogenesis in Mucor circinelloides
 Thesis (2001) Department of Microbiologia y Genetica, Universidad
 de Salamanca, Salamanca, Spain
 2
 Velavos, A., Fuentes, M., Aguilar, R., Esilava, A.P. and Iturriaga, E.A.
 Prenyl synthases in Mucor circinelloides
 Unpublished
 3 (bases 1 to 2666)
 Velavos, A.
 Direct Submission
 Submitted (22-JUL-2002) Velavos A., Microbiologia y Genetica, Area
 Genetica, Facultad de Biologia, Universidad de Salamanca, Edificio
 Departamental, Avenida del Campo Charro s/n, Salamanca, E-37007,
 SPAIN
 Location/Qualifiers
 1. .2666
 /organism="Mucor circinelloides f. lusitanicus"
 /strain="CBS277.49"
 /db_xref="taxon:29924"
 /note="forma: lusitanicus"
 515. .2091
 /gene="isob"
 join(515..1594,1650..1823,1893..2054)
 /gene="isob"
 /EC_number="2.5.1.11"
 /codon_start=1
 FEATURES
 source
 gene
 CDS

```

/product="solanesyl pyrophosphate synthase"
/protein_id="CAD42868.1"
/db_xref="GI:21955862"
/translation="MFVVRQSKALSPVIRHRSIAASFPTTAMLATPAKESDNDASAR
EKIHGNSFKSTVSGSPRLNTRKPVPTPTGSEANNOALSEQSLVKSDRRIID
PAKIVGDLKREIKANISKLSSGHPFLNTVARRHFTSGSKVRLVLLIQAQNSIAD
KKSLSANIDYKSIDTPISHGLKNTITNDVFRKLHTYTSVTDGCTILPTORLAEIS
EMIHATSLIHDVDIDASWTRNRLPSANASFNKQAVLGLDPLLAASLALARNAC
ELMATHSLANVEGFEMQLRNKEGSEKVKLTFDHYMEKTYAKTSLIAQSKAS
AVLGSTKEVANIAYDFGKNLGLAFQVLVDMDIDFTVAELKGPAGADLKGLATAPV
LFAMEYREPELPIKRKFSVKDEKARDLVYSGDLKKTLIDLAQHCKLATDMLVKL
PASDARSALVOITETKLTRRS"
exon
515..1594
/gene="1594"
/intron
1595..1649
/gene="1594"
exon
1650..1823
/gene="1594"
intron
1824..1892
/gene="1594"
exon
1893..2054
/gene="1594"
polyA_signal
2054..2059
/gene="1594"
polyA_site
2091
/gene="1594"
BASE COUNT 741 a 692 c 538 g 695 t
ORIGIN
Query Match 14.0%; Score 231.4; DB 8; Length 2666;
Best Local Similarity 54.6%; Pred. No. 1.9e-39;
Matches 556; Conservative 0; Mismatches 396; Indels 67; Gaps 2;
QY 327 TCATCAATCCGCTCAACTTGTGTCGTCGCCAGATGCAATCTTATCATCAACATCCGAT 386
Db 813 TAAATGATCTCTGCCAACTAGTGGAGAGATCTCAAGAAATCAAGGCAACATCTTCA 872
QY 387 CTCTCTCGGTTCAAGACACCTTCTCTGACACCTGCTTAATCTATGTTCAAGTCTG 446
Db 873 AATTGCTGGGCGAGTGGACACCCGTTCTCAACACCGTGGCCAGACACTTATAGCGCG 932
QY 447 AGGAAAGCATTTTGTCCGCTCATAGTACTGCTGATGCTCAGGAGGAGGAGTTGCC 506
Db 933 AGGTAAAGCATGTGCAACCATGCTGTGCTGTATAGCGGACGACACCATGCGAG 992
QY 507 CAAAAGTTCAGGGTTGGAGAGTCTGTGAGAGTTCCGGTGAACGAGGAGCTCGACAC 566
Db 993 ACMAA-----AAGTGGCTGGCCAGCAACCTTATGACATCAAAATGATGACAC 1040
QY 567 CAGAGTGTCTAATGACAAAGAACCAATATATGATGAACATGAGTCAAGACATTAACA 626
Db 1041 CTATATCGATGAGCTCAAGAAACATCAACCAACGACCTGTTGACCGCAAGCTTACT 1100
QY 627 AGGAGGGGAGATTCAGGAGACAGAGCTGCAATATCTCGCTCGCAACGGCGGTTGGCT 686
Db 1101 ACACACCCCTCCGTGACAGATGAGGATGATCATCTTACCCACACGACGACGATTAAGC 1160
QY 687 AGATACGGAGATGATTCATGAGCATCTCTCCAGACGACGTTATCGAGCTTCCG 746
Db 1161 AATATTCGAGATGATTCATGAGGCTCTTATACACATGATGATGATGAGCGCTTCA 1220
QY 747 AGACGAGAGAAACGACCATCCGAAACAGGCAATTCGAAACAAAGATGGGATTTTG 806
Db 1221 TGACTCGGAGAACTTCCGCTCGCAATGCGTGTGTTGGCAACAAATGAGGAGTCTCG 1280
QY 807 CTGTGATTTTCTTTGGAGCGGCTGTGTTGATTTGGCGAGTTGCCAATCCGAGAG 866
Db 1281 GCGGAGCATTTCTTGTGGGCAAGGATTCATTTGGCAAGACTCAAGAAACGAGAGT 1340

```

```

QY 867 TGATTGACCTTTGGCTACTGTTATTCGAACTTGTTAGGAGAGTTCAAGCTTGA 926
Db 1341 GCATGGAATTAATGAGCCACATGATATCGCAATCTGTGGAGGCGCAATTCATGCAATGA 1400
QY 927 AAAATCTGTTGATGATTCGATTAAGGCTAAGCGGACGAGAAACGTTGATTAATT 986
Db 1401 GAAACACCAAGAGAGGAATTCGAGCAAGTGAAGAAACCTGAGCACTTTGACATTAACA 1460
QY 987 TGCAGAACTTACTTGAAGACTGCGTCTTATGTTGCCAAGTCGACAGCAAGTGGC 1046
Db 1461 TCGAGAAACCTTACATGAAGACGCGCACTGATTTGCCAAGCTTCAAGCGCTGCGC 1520
QY 1047 TTCTGGTGTGCTTACGCTTGAAGTGTGCTGATGCTTATGCTTATTCGAAAGAACTTG 1106
Db 1521 TACTGGTGTGCTCCACCAAGAAAGTCCCAATTCATATGATTTTGAAGAAATCTGG 1580
QY 1107 GTTGGCATTC----- 1118
Db 1581 GTTAGCCTTTCAAGTAAATACACAGCAATATCTTATCAACCCCTGTAACACTTTTC 1640
QY 1119 -----AGATGTCGACGACATGCTGACTACACCGTCTCGCTACCGACTGGTAAG 1171
Db 1641 TCTGATAGCTTTGATGATGATGATGCTGATTTACAGAGACAGCAGAACTGGGCAAG 1700
QY 1172 CCGGCGGTGACAGACCTTCAGCTGCTGCTGCAACCGCGCGCGCTTCTGCAATGAG 1231
Db 1701 CCGGAGAGGAGGATCTCAAGTGGGCTGGCCACAGCGCCGCTGTGTTGTTGCTTGGGAG 1760
QY 1232 CACCAAGCGGAGCTGCTGCTGCTGATGATCAAGGCAAGTTCTTGACCCAGAGAGCTGCA 1290
Db 1761 GAGTATCCAGATTTGAGACCATTTGATCAAGGAAAGTTCTGTAAAGGCGATGAGGA 1819

RESULT 4
E13333
LOCUS E13333 1134 bp DNA linear PAT 27-APR-1998
DEFINITION gDNA encoding prenyl diphosphate synthase.
ACCESSION E13333
VERSION E13333.1 GI:3252138
KEYWORDS JP 1997173076-A/1.
SOURCE Schizosaccharomyces pombe.
ORGANISM Schizosaccharomyces pombe
REFERENCE 1 (bases 1 to 1134)
AUTHORS Matsuda,H., Kawamuki,M. and Nakagawa,T.
TITLE PRODUCTION OF DIFFERENT KIND UBIOUINONE
PATENT: JP 1997173076-A 1 08-JUL-1997;
JOURNAL ALPHA-SHOKUHIN KK
COMMENT OS Schizosaccharomyces pombe
PN JP 1997173076-A/1
PD 08-JUL-1997
PE 27-DEC-1995 JP 1995351243
PI MATSUDA HIDEYUKI, KAWAMUKI MAKOTO, NAKAGAWA TSUYOSHI PC
C12N15/09,C12N9/10,C12P7/66,C12N15/09,C12R1/19),(C12N15/09,PC
C12R1:645),
PC (C12N9/10,C12R1:865);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH key Location/Qualifiers
FT source 1..1134
FT FT /product='Prenyl diphosphate synthase'.
FEATURES
source location/Qualifiers
1..1134 /organism='Schizosaccharomyces pombe'
/db_xref='taxon:4896'

```


BASE COUNT 328 a 221 c 243 g 342 t
 ORIGIN
 Query Match 13.4%; Score 222; DB 6; Length 1134;
 Best Local Similarity 56.6%; Pred. No. 2e-37;
 Matches 452; Conservative 0; Mismatches 340; Indels 6; Gaps 2;
 QY 649 GAGTGCATATCTCGCTCGCAACGGCGTTGGCTGAGATCACGGAGATGATCCATGC 708
 Db 333 GACGGTCAAAATCTTCCTCTCAATGAGATTAGACAAATAACCGAGATGATCCATAT 392
 QY 709 AGCATCACTCCTCCACAGACAGCTTATCGACGCTTCGAGACGAGCAAAACGACCATC 768
 Db 393 AGCAAGTTTCTGCATGACAGATGATGATACGCTAATGTCCTAGAGGCTCACCTTC 452
 QY 769 CGAAACACGAGCATTCGGAACAAAGATGCGATTTTGGCTGGTGAATTTCTTGTGGAGC 828
 Db 453 AAGCAATGTTGCTTTCGGTAAATCGACGGTCAATCCTTGGGGTAAATTCATCCTTCACG 512
 QY 829 GCGTCTGTGCTATGCGAGGTTGGCAATCCGGAGGTGATTGAGCTTTTGGCTACTGT 888
 Db 513 GCGTTCGACTGTATGCGCCGCTTCGAAATCCCAAGTTACGGAGTTGTTAGCTACAGT 572
 QY 889 TATTGCAAACTTGGTTGAGGAGAGTTTCATGCAAGTTGAAATAACTGTTGATGATCGAT 948
 Db 573 GATAGCAGACTTGGTTGAGGTGAGTTTTCGAGCTTAAATAATACTATGATCCTTCATC 632
 QY 949 TGAGGCTACGGGACGACGAGAAACGTTTCGATTAATTTTCGAGAAGACTTACTTGAAGAC 1008
 Db 633 TTTGGAAA---TAAACAATCAAAATTTGACTATTATATGAAAAAGTTTITGAA--- 686
 QY 1009 TGGTCTCTGATTGCCAAGTCGTGCGAGCAAGTGGCTTCTGGGTGGTGTACGCGCTGA 1068
 Db 687 ACAGCAGATTATTTTCAAAAGCTGCAAGGCTTCTACAATCCTCGACAATGTTCTCTAC 746
 QY 1069 GGTGCTGATGCTGTATGCTTACGGAAGAACCTTGGTTTGGCATTCAGATTCCTCGA 1128
 Db 747 TGAGCAACAGCTGCTGGGGAATACGGTCGATGCTATGCTGCTTTTCAACTAATGGA 806
 QY 1129 CGACATGCTCGACTACACCGCTCTCCGCTACCGACCTCGGTAAAGCCCGGTCGACACCT 1188
 Db 807 TGACGTTTGGACTATACGTCGAAAGATGATACTTTAGGAAGCGCGTGTGTCAGATT 866
 QY 1189 CMAGTCCGTTCTCCACCGCGCGCCCTCTTCGATGAAAGCACACGCGAGTCGG 1248
 Db 867 GAAGCTAGGGTTGGCTACAGCTCCGCTCTTTCGATGAAAAAGTATCCAGAACTTGG 926
 QY 1249 TCCATGATCAACGCGACGTTCTGACCCAGGAGCGTTCGACGCTGCAGCGAGTTGGT 1308
 Db 927 TGAATGATTGTAATAGATTCAATCATCTTCTGATATCCAACGGGCTCGTTCTTTGGT 986
 QY 1309 CGAGAAAAAGTATGGATTGGAGAAAGACGAGAGCTTGGCGGAGGAGTATGCCAGAAAGC 1368
 Db 987 TGAGTCACTGATGCTATCGAGCAACCATCATCTTGGCAAAAGATATATCAAAAAGC 1046
 QY 1369 GTTGGATGCAATTCGACGTTCCCGAGAGTCCGCGACGGAAGCTTTGGAGCAGTTGAC 1428
 Db 1047 CAAAGATTCCCTTCTGTCTCCTGATTACCTGATTCACCTGCAAGGAAGGCACATTTTTCGGTTGC 1106
 QY 1429 GGACAGGTGTTCACTAG 1446
 Db 1107 TGATAAGTAATAACGAG 1124

RESULT 5

D89265
 LOCUS D89265 1492 bp mRNA linear PLN 13-MAR-1998
 DEFINITION Schizosaccharomyces pombe mRNA, partial cds, clone: SY 1754.
 ACCESSION D89265
 VERSION D89265.1 GI:1749737
 KEYWORDS Schizosaccharomyces pombe (strain:PR745) cDNA to mRNA,
 SOURCE clone_lib:library of H. Nojima clone:SY1754.

Schizosaccharomyces pombe
 Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 Schizosaccharomycetales; Schizosaccharomycetaceae;
 Schizosaccharomycetes.
 1 (sites)
 Yoshioka,S., Kato,K., Nakai,K., Okayama,H. and Nojima,H.
 Identification of open reading frames in Schizosaccharomyces pombe
 CDNAS
 DNA Res. 4 (6), 363-369 (1997)
 98162722
 2 (bases 1 to 1492)
 Yoshioka,S.
 Direct Submission
 Submitted (15-NOV-1996) Sachiyo Yoshioka, Tsukita Cell Axis Project
 ERATO JST, Kyoto Research Park; 17 Chudouji Minamimachi,
 Shimokyo-Ku, Kyoto 600, Japan
 (E-mail: syoshi@cell.tsukita.jst.go.jp. Tel.: +81-75-315-7913,
 Fax: +81-75-315-6420)
 Location/Qualifiers
 1. .1492
 /organism="Schizosaccharomyces pombe"
 /strain="PR745"
 /db_xref="taxon:4896"
 /clone="SY1754"
 /clone_lib="library of H. Nojima"
 40. .1110
 /note="similar to Saccharomyces cerevisiae hexaphenyl
 pyrophosphate synthetase precursor, SWISS-PROT Accession
 Number p18900"
 /codon_start=1
 /protein_id="BAA13926.1"
 /db_xref="GI:1749738"
 /translation="MIQVYIKHMKRLMSLKVRSVTLRFTNNRASHLIKNELEQI
 SPGIQMLNSSEFLIEGSKYVTAQGMKRPISLVLMKSTLSLPHGIDRSVVGDKYT
 DDDLRFSFTGILPSLAQITEMIHLASLLHDDVIDHANVRGSPSSNVAFGNRR
 SILAGNFILARASTVMARLNPOVTELLATVIADLRGEFLQKNTMDPSSLEIKQSN
 FDYIEKSLFKTASLISKSKASTILGQCSPTVAAGEYGRICGTAFQMLDDVLDYT
 SKDPTLGAAGADLKLGLATAPLVFAWKYPKELGAMIVNRFNHPDSIQARSLVBCSD
 AIEQDPTWAIEYIKRFPSPSP"
 BASE COUNT 453 a 282 c 291 g 466 t
 ORIGIN
 Query Match 13.2%; Score 218.6; DB 8; Length 1492;
 Best Local Similarity 56.5%; Pred. No. 1.1e-36;
 Matches 450; Conservative 0; Mismatches 339; Indels 8; Gaps 2;
 QY 650 AGTTCGAATATCTCGCTCGCAACGGCGTTGGCTGAGATCACGGAGATGATCCATGCA 709
 Db 373 ACGGTCAAAATCTTCCCTTCAATTGAGATTAGCAAAATACCGAGATGATCCATATA 432
 QY 710 GCATCACTCTCCACGACGAGCTTATCGACGCTTCGAGACCAAGCAAAACGACCATCC 769
 Db 433 GCNAGTTTCTGCATGACGATGTTGATTGATCAGCTAATGTCGTAGAGGCTCACCTTCA 492
 QY 770 GGAACACGAGCATTCGGAACAAAGATGGCGATTTTGGCTGGTGAATTTCTTGTGGACGG 829
 Db 493 AGCAATGTTGCTTTCGGTAAATCGACGCTCAATCCTTTCGCGGTAATTTCACTCTTGCACGG 552
 QY 830 GCGTCTGTTGCTATGCGAGGTTGCCAATCCGAGGTGATTGAGCTTTTGGCTACTGTT 889
 Db 553 GCTTCGACTGTTATGGCCCGCTTCGAAATCCCAAGTTCGAGAGTTGTTAGTACAGTG 612
 QY 890 ATTGCAAACTTGGTTGAGGAGAGTTTCATGCAAGTTCAAAAATACTGTTGATGATGGGATT 949
 Db 613 ATAGCAGACTTGGTTCGCGGTGAGTTTTCAGCTAAAAATACTATGATCCTTCTACTCT 672
 QY 950 GAGGCTACGGCGACGCGAGAAACGTTTCGATTACTATTTCGAGAAGACTTACTTTGAAGACT 1009
 Db 673 TTGGAAA---TAAACAATCAAAATTTGACTATTATATTGAAAAAAGTTTTTTTGAACA 729
 QY 1010 GCGTCTCTGATTGCCAAGTCGTCGACGAGCAAGTCCGCTTCTGGGTGGTGTAGCCTGAG 1069
 Db 730 GCCAGTTTAAATTTCCAAAAGCTGCAAGGCTTCTACAATCCTCGGACAATGTTCTCTCTACT 789

QY 1070 GTTGTGATGCTGCTTATGCTTACGGAAGAACTTGGCATTTCCAGATGCTGCAC 1129
 Db 790 GTAGCAACAGCTGTGAGAAATACGCTCATGATGCTGCTTTTCACTTAAGAT 849
 QY 1130 GACATGCTGACTTACACCGCTCCGCTACCGAAGCTGAGAGCCGCGTGCAGACTC 1189
 Db 850 GACGTTGGACTTATACGTCGAAAGATGATCTTATGAGAAAGCGCTGCTGAGATTGG 909
 QY 1190 CAGCTGCTGCTGCGCAACCGCGCGGCTCTTTCGATGAGAACACACCGCAAGCTCGT 1249
 Db 910 AAGCTAGGGTGGCTACAGACTCCGCTCTTTGATGAGAAAAGATATCAGAACTGGT 969
 QY 1250 CCCATGATCAACGCGAAGTCTTCTGACGAGAGAGCTGAGCGCTGCACGCGAGTTGCTC 1309
 Db 970 GCAATGATTTGTAATAGATGATCATCTCTGATATCAACGCGGCTGCTTCTGCTG 1029
 QY 1310 GAGAAAGATGATGATGAGAAAGAGAGAGCTTGGCGGAGAGATGCCAGAAAGCG 1369
 Db 1030 GAGTCTCTGATGCTATCGAGAACCCATCATCTTGGGCAATGATATATCAAAAA---- 1085
 QY 1370 TTGATGCAATTCGAGACGTTCCCGGAGAGTCCGCGACGAGAGCTTTGGAGCAGTTGACG 1429
 Db 1086 -AAGATTCCTTCTGCTGCTCCCTGATTCACCTCGAAGAGGCACTTTTGGCTGGCT 1144
 QY 1430 GACAAAGTGTGACTAG 1446
 Db 1145 GATAAGTATATACGAG 1161

RESULT 6
 SPBJ694 2097 bp DNA linear PLN 24-APR-2001
 LOCUS S.pombe chromosome II PCR product pJ694.
 DEFINITION AL136498
 ACCESSION AL136498.2 GI:13810218
 VERSION decaprenyl diphosphate synthase; dps.
 KEYWORDS fission yeast.
 SOURCE Schizosaccharomyces pombe
 ORGANISM Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 Schizosaccharomycetales; Schizosaccharomycetaceae;
 Schizosaccharomyces.
 1 (bases 1 to 2097)
 McDougall, R.C., Rajandream, M.A., Barrell, B.G., Saunders, D. and
 Harris, D.
 TITLE Direct Submission
 JOURNAL Submitted (17-JAN-2000) European Schizosaccharomyces genome
 sequencing project, Sanger Centre, The Wellcome Trust Genome
 Campus, Hinxton, Cambridge CB10 1SA, E-mail: Barrell@sanger.ac.uk
 On Apr 26, 2001 this sequence version replaced gi:16714818.
 COMMENT Notes:
 Details of yeast sequencing at the Sanger Centre are available on
 the World Wide Web.
 (URL, http://www.sanger.ac.uk/projects/S_pombe/)
 During 1995 to 1996 about 66% of S. pombe chromosome 1 was
 sequenced by the Sanger Centre. The sequencing of the S. pombe
 genome is now being continued with funding from The European
 Commission. Fourteen European sequencing laboratories, including
 the Sanger Centre, are participating in the project.
 Protein coding regions (CDS) have been predicted with the help of
 computer analysis using the GeneIndex program in Pombase (an ACEDB
 database) with additional predictions for the branch-acceptor sites
 supplied by the program Splice. CAUTION: It is possible that for
 any individual CDS we may have underestimated or overestimated the
 number of introns/exons or we may not have chosen the correct
 splice donor/acceptor sites. CDS are numbered using the following
 system eg SPBC25H2.01c. SP (S. pombe), B (chromosome 2), c25H2
 (cosmid name), .01 (first CDS), c (complementary strand).
 The more significant matches with motifs in the POSING database
 are also included but some of these may be fortuitous. The length
 in codons is given for each CDS.
 IMPORTANT: This sequence MAY NOT be the entire insert of the
 sequenced clone. It may be shorter because we only sequence

overlapping sections once, or longer, because we arrange for a
 small overlap between neighbouring submissions. PCR product pJ694
 is overlapped by cosmid c947, EMBL entry SPBC947, accession number
 AL021837, at the 5' end.

FEATURES
 source
 1..2097
 /organism="Schizosaccharomyces pombe"
 /strain="972h-"
 /db_xref="taxon:4896"
 /chromosome="II"
 /map="11L"
 1..100
 /note="nominal overlap with cosmid SPBC947, EM:AL021837 S.
 pombe chromosome 2"
 join(856..1830,1912..2073)
 /gene="dps"
 /note="SPBJ694.01; SPBJ4664.01"
 join(856..1830,1912..2073)
 /gene="SPBJ4664.01"
 /note="SPBJ694.01, len:378"
 /codon_start=1
 /label=dps
 /product="decaprenyl diphosphate synthase"
 /protein_id="CA86154.1"
 /db_xref="GI:6714819"
 /db_xref="SPRMBL:O43091"
 /translation="MIQVYVLKRMKLSLQKVRSTVLRSTTRNASHLIRNELQI
 SPGRIMLNSEFLIECSKXYITIAQCKORPSVLVLSKATSLCHIDISVGDXYI
 DDDLRFSFTGQILPSQRLAQIEMHISLIDVDYIDHANVRSQSSNVAGNR
 SILAGNFIARASTAMARLRNPOTVELATVLDVNGEFLQKNTMPSLEIKQSN
 PDYIERSPLKTSALISKSKASITLQCCPYATAAGEGRICGTAFQMDVDVDT
 SKDITLRAGADIKLGLATAPVLFPAKKIPELGAMIVNRPNSDIQRARSLEVEDT
 AIEQITVMAKEYIKKADSLCLPDSPARALPALADKVTTRKK"
 join(1030..1100,1179..1866)
 /gene="dps"
 /note="match to PF00348 polyprenyl_synth, Polyprenyl
 synthetases Score 137.19"
 1831..1911
 /gene="dps"
 /note="intron confirmed by EST"
 1831..1836
 /gene="dps"
 /note="gtaga, splice donor sequence"
 1896..1911
 /gene="dps"
 /note="ctacatacatctag, splice branch and acceptor"
 BASE COUNT 682 a 355 c 383 g 677 t
 ORIGIN
 Query Match 12.1%; Score 199.6; DB 8; Length 2097;
 Best Local Similarity 57.6%; Pred. No. 1.3e-32;
 Matches 377; Conservative 0; Mismatches 274; Indels 3; Gaps 1;
 QY 649 GACGTCGAATATCCCGCTCGCAACGCGGCTGAGATCAACGAGATGATCCAGC 708
 Db 1188 GACGGGTCAAAATCTTCTTCAATTAGATGACAAATACCGAGATGATCAAT 1247
 QY 709 AGCATCACTCTCCACGACGAGATGATGACGCTCCGAGACGACAAACGACCATC 768
 Db 1248 AGCAAGTTGCTGCAATGACATGATGATGATGATGATGATGATGATGATGATGAT 1307
 QY 769 CGAAACGAGGATTCGAAACAGAGTGGCATTTTGGCTGATTTCTTGTGGAGC 828
 Db 1308 AAGCAATGTTGGCTTCGGTAAATCGACGCTCAATCTTCGGGTATTTTCACTTCGACG 1367
 QY 829 GGCCTCTGTTGATGAGGAGGATGCGCAATCCGAGATGATGATGATGATGATGATGAT 888
 Db 1368 GCTTCGACTGTATGAGCGCGCTTCGAAATCCCAAGTTACGAGTTGATGATGATGAT 1427
 QY 889 TATTCGAATCTGTTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
 Db 1428 GATGACAGCTTTGTTTCAGAGTGAATTTTTCGAGCTATAAAAAATGATGATGATGATGAT 1487

```

FEATURES             source
1. .5371
   Location/Qualifiers
     /organism="Cloning vector pBAD-GFPuv"
     /db_xref="taxon:50707"
     complement(96..974)
     /gene="araC"
     complement(96..974)
     /gene="araC"
     /note="PID: g455167"
     /codon_start=1

```

/transl_table=11	
/product="araC protein"	
/protein_id="AAC53662.1"	
/db_xref="GI:1490532"	
/translation="MAEAQNDELLPGYSNAHLVAGLTPIEANGYLDFFIDRPLGNKGG	
YIUNLIRIQGVVKNQREFVKPGDILLFPPEGIHHYGRHPPEARWYHQWVYFRPRA	
YHWEMLNWSIFANTGFFRDEAHQPHFSLDFGOIINAGQEGREYSSELLAINLLEQL	
LRMEAINESLHPMDNRVREACQYISDLADSNFDIASVAQHVCLSPRLSHLFRQQ	
LGI SVLSWREDQRIISOAKLLLS"TRMPIATVGRNVGFDQQLYPSRVFKKCTGTGASPSF	
RACCEBKVNDVAVKLS"	
gene	1342. .2061
/gene="gfpuv"	
CDS	1342. .2061
/gene="gfpuv"	
/notes="GFPuv is the GFP variant called 'cycle 3'; Allele:	
AC2: green fluorescent protein variant"	
/codon_start=1	
/transl_table=11	
/product="GFPuv"	
/protein_id="AAC53663.1"	
/db_xref="GI:1490533"	
/translation="MASEGELFTGVVPILVELDGVNKGHSVSGEGEDATYGLKT	
LKEPCTIGKLPVWPPLVITVTSYGVQCFSRYDPDMKRRHDFFKSAMPEGVQERTISFK	
DGNYKTRAEVKEGDTLWRIELKGIDFKEDGNLGHKLEYNINSHNVIITADKOKN	
GIGANKFKIRHNEDTSGVQLADHYQQNTPIGDGVPLLPDHNHLYSTQSALSKDPNKEKRDH	
MVLLFVTAAGITGHGMDLYK"	
gene	2636. .3496
/gene="bla"	
CDS	2636. .3496
/gene="bla"	
/function="confers resistance to ampicillin"	
/codon_start=1	
/transl_table=11	
/product="beta-lactamase"	
/protein_id="AAC53664.1"	
/db_xref="GI:1490534"	
/translation="MSIQHFRVALIPFFAAFCIPVFAHPETLVKVKDAEDQLGARVGY	
IEDLNSKGILLESFPEERFPMWSTFKVLICGAVLSRDVAGQOLGRRIRHYSDNDAVE	
YSPVTEKHETDGGTVRELCSAALTMSDNTAANLLTTIGPKELTAFLHMGDHTRL	
DRWEPELNAIPNDERDTPWAMATTIRKLITGELLTLASRQQLIDWMEADKVKAGPL	
LRSALPAGWFIADKSGABGRSGRIIAALGPDGKPSRIVVIYTTGSAQTMDERNRQIA	
ETGASLIKHW"	
BASE COUNT	1369 a 1368 c 1300 g 1334 t
ORIGIN	
Query Match	12.0%; Score 198.6; DB 12; Length 5371;
Best Local Similarity	98.0%; Pred. No. 2.2e-32;
Matches	201; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1449	CAAGATAGGAATTCGACCTCGGTACCGCGGGATCCTCTAGACTGCACCTCAGCGATGCA 1508
Db 2055	CAATAATGAATTCGACCTCGGTACCGCGGGATCCTCTAGACTGCACCTCAGCGATGCA 2114
Qy 1509	AGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTTCAGCCTGATACAGATTAATAATCAGAAC 1568
Db 2115	AGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAATAATCAGAAC 2174
Qy 1569	GCAGAAGCGGTCTGATAAAAACAGATTTCCTCGCGGCAGTAGCGCGGTGGTCCCACTG 1628
Db 2175	GCAGAAGCGGTCTGATAAAAACAGATTTCCTCGCGGCAGTAGCGCGGTGGTCCCACTG 2234
Qy 1629	ACCCCATGCCGAACTCAGAAGTGA 1653
Db 2235	ACCCCATGCCGAACTCAGAAGTGA 2259
RESULT 8	
Al3038	
LOCUS	Al3038 4176 bp DNA linear PAT 03-JAN-1994
DEFINITION	Artificial sequence of plasmid pTrc99A.
ACCESSION	Al3038
VERSION	Al3038.1 GI:491502
KEYWORDS	

SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 4176)
AUTHORS Amann, E. and Abel, K.J.
TITLE Expression vectors for the production of non-fusion proteins in microorganisms
JOURNAL Patent: EP 0345615-A 5 13-DEC-1989;
BEHRINGERWERKE Aktiengesellschaft
FEATURES
source 1..4176
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 1009 a 1059 c 1118 g 990 t
ORIGIN
Query Match 12.0%; Score 198; DB 6; Length 4176;
Best Local Similarity 100.0%; Pred. No. 3e-32;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1456 GGAATTCGAGCTCGGTACCGCGGGATCCTTAGAGTCGACCTGCAGGCGATGCAAGCTTGG 1515
Db 269 GGAATTCGAGCTCGGTACCGCGGGATCCTTAGAGTCGACCTGCAGGCGATGCAAGCTTGG 328
QY 1516 CTGTTTGGCGGATGAGAGAGATTTTCAGCTGATACGATTAAATCGAACCGAGAAG 1575
Db 329 CTGTTTGGCGGATGAGAGAGATTTTCAGCTGATACGATTAAATCGAACCGAGAAG 388
QY 1576 CGGTCTGATTAACAGAAATTTGCTGGCGGAGTAGCGGAGTGTCCACCTGACCCCAT 1635
Db 389 CGGTCTGATTAACAGAAATTTGCTGGCGGAGTAGCGGAGTGTCCACCTGACCCCAT 448
QY 1636 GCCGAACCTCAGAAAGTGAA 1653
Db 449 GCCGAACCTCAGAAAGTGAA 466
RESULT 9
LOCUS SYNTRC99A 4176 bp DNA linear SYN 29-JUN-1995
DEFINITION Cloning vector pTRC99A, complete sequence.
ACCESSION M22744
VERSION M22744.1 GI:340740
KEYWORDS cloning vector.
SOURCE Cloning vector pTRC99A.
ORGANISM Cloning vector pTRC99A.
REFERENCE 1 (sites)
AUTHORS Amann, E. and Brosius, J.
TITLE 'ATG vectors' for regulated high-level expression of cloned genes in Escherichia coli
JOURNAL Gene 40 (2-3), 183-190 (1985)
MEDLINE 86165860
PUBMED 3007288
REFERENCE 2 (bases 1 to 4176)
AUTHORS Amann, E., Ochs, B. and Abel, K.J.
TITLE tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in Escherichia coli
JOURNAL Gene 69 (2), 301-315 (1988)
MEDLINE 8912074
PUBMED 3069586
FEATURES
source Location/Qualifiers
1..4176
/organism="Cloning vector pTRC99A"
/db_xref="taxon:40992"
/lab_host="Escherichia coli"
/note="derived from pKK233-2"
1..17
/note="derived from cloning vector pBR322"
18..263
/note="tac promoter from pKK233-2"
/citation=[1]
misc_feature 264..270

misc_feature /note="NcoI/EcoRI linker"
271..321
/note="derived from cloning vector pUC18"
misc_feature 329..753
/note="5S RNA, T1, T2, rRNA"
misc_feature 754..2932
/note="derived from cloning vector pBR322"
misc_feature 2933..2940
/note="BgIII linker"
misc_feature 2933..2938
/note="BglII linker"
misc_feature 2939..2945
/note="EcoRI linker"
misc_feature 2946..2976
/note="derived from plasmid RP4"
misc_feature 2977..4163
/note="lacI-q region"
misc_feature 4164..4170
/note="EcoRI linker"
misc_feature 4171..4176
/note="BglIII linker"
BASE COUNT 1009 a 1059 c 1118 g 990 t
ORIGIN
Query Match 12.0%; Score 198; DB 12; Length 4176;
Best Local Similarity 100.0%; Pred. No. 3e-32;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1456 GGAATTCGAGCTCGGTACCGCGGGATCCTTAGAGTCGACCTGCAGGCGATGCAAGCTTGG 1515
Db 269 GGAATTCGAGCTCGGTACCGCGGGATCCTTAGAGTCGACCTGCAGGCGATGCAAGCTTGG 328
QY 1516 CTGTTTGGCGGATGAGAGAGATTTTCAGCTGATACGATTAAATCGAACCGAGAAG 1575
Db 329 CTGTTTGGCGGATGAGAGAGATTTTCAGCTGATACGATTAAATCGAACCGAGAAG 388
QY 1576 CGGTCTGATTAACAGAAATTTGCTGGCGGAGTAGCGGAGTGTCCACCTGACCCCAT 1635
Db 389 CGGTCTGATTAACAGAAATTTGCTGGCGGAGTAGCGGAGTGTCCACCTGACCCCAT 448
QY 1636 GCCGAACCTCAGAAAGTGAA 1653
Db 449 GCCGAACCTCAGAAAGTGAA 466
RESULT 10
LOCUS XXU13872 4176 bp DNA circular SYN 07-DEC-1994
DEFINITION pTRC99A cloning vector, complete sequence.
ACCESSION U13872
VERSION U13872.1 GI:595782
KEYWORDS 5S; beta-lactamase; lac repressor protein.
SOURCE unidentified cloning vector.
ORGANISM unidentified cloning vector.
REFERENCE 1 (bases 1 to 4176)
AUTHORS Malone, J.A.
TITLE pTRC99A: A cloning vector for inducible expression of cloned inserts using the trc promoter
JOURNAL unpublished (1994)
REFERENCE 2 (bases 1 to 4176)
AUTHORS Amann, E., Ochs, B. and Abel, K.J.
TITLE tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in Escherichia coli
JOURNAL Gene 69 (2), 301-315 (1988)
MEDLINE 8912074
PUBMED 3069586
REFERENCE 3 (bases 193 to 222)
AUTHORS Brosius, J., Bfille, M. and Storella, J.
TITLE Spacing of the -10 and -35 regions in the tac promoter. Effect on its in vivo activity
JOURNAL J. Biol. Chem. 260 (6), 3539-3541 (1985)
MEDLINE 85131154

```

PUBMED 2579077
REFERENCE 4 (bases 328 to 753)
AUTHORS Brosius,J., Dull,T.J., Sleeter,D.D. and Noller,H.F.
TITLE Gene organization and primary structure of a ribosomal RNA operon
from Escherichia coli
J. Mol. Biol. 148 (2), 107-127 (1981)
MEDLINE 8205570
PUBMED 7028991
REFERENCE 5 (bases 1 to 4176)
AUTHORS Malone,J.A.
TITLE Direct Submission
JOURNAL Submitted (19-AUG-1994) James A. Malone, International Technical
Services, Molecular Biology Reagents Division, Pharmacia Biotech
Inc., 2202 N. Bartlett Ave., Milwaukee, WI 53202-1009, USA
FEATURES
source
1..4176
/organism="unidentified cloning vector"
/db_xref="taxon:45196"
/lab_host="Escherichia coli"
promoter
133..222
/notes="The trc promoter combines the -35 region from the
trp promoter and the -10 region from the lacUV5 promoter"
/citation=[3]
193..198
/genes="trp"
-35_signal
133..198
/genes="trp"
gene
216..222
/genes="lacUV5"
-10_signal
216..222
/genes="lacUV5"
protein_bind
228..248
/bound_moiety="lacIq repressor protein"
misc_feature
265..314
/notes="Multiple Cloning Site (MCS); contains unique sites
for NcoI, EcoRI, SacI, KpnI, SmaI, BamHI, XbaI, SalI,
PstI, and HindIII"
328..753
/notes="derived from the rrnB operon; the entire operon is
not contained in this vector"
409..734
/genes="rrnB"
rRNA
409..528
/genes="rrnB"
terminator
532..575
/product="5S ribosomal RNA"
/genes="rrnB"
/standard_name="rrnB T1"
terminator
707..734
/genes="rrnB"
/standard_name="rrnB T2"
gene
776..1706
/genes="bla"
promoter
776..804
/genes="bla"
CDS
846..1706
/genes="bla"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AA56743.1"
/db_xref="GI:595783"
/translation="MSIQHFRVALIPFFAFLPVFAHPETLVKVKDAEDQLGARVGY
IEDLNKSGILSPERFPMWSTFKLLCGAVLSRVDAVGQOLGRRTHYQNDLVE
YSPVTEKHLTDGMTVELCSAALTMSDNTAANLLLTIGGPKELTAFLLHMGDHYTRL
DWEPELNEAIPIRDERDTPMTAMATTLKLTIGELLTLASQQILIDWHEADKVAQPL
LKSALPAGFWFIADKSGAGRGRIIAGLPKPSRIIVYITGSGATMDERNQIIA
EIGASLIKHV"
2464
/notes="base 2464 represents the first base of the newly
synthesized single strand"
/direction=right
3055..4137
gene

```

```

CDS
/genes="lacIq"
3055..4137
/genes="lacIq"
/codon_start=1
/transl_table=11
/product="lac repressor protein"
/protein_id="AA56744.1"
/db_xref="GI:595784"
/translation="MKPVTLYDVAEYAGVSQTVSRVNVQASHVSAKTRKVEAAWAE
LNYIPNRVAQQLAGKQSLIGVATSSLAHAPSQIVAAIKSRADQLDGASVVVSMVRS
GYEACKAAVHNLIAQRVSGLIINYPDQDAIAVEAACTNVNPAFLDVSQDQPIINSII
FSHEGTGLGVHILVALGHOQIALLAGPLSSVSRRLAGWHKYLTRNOIOPIAREEG
DHSAMSGQQIMQMLNEGIVPTAMLVANDQALGAMRAITESGLVAGDISVVGVDIT
EDSSCIPIPTTIKODFRLLGGTSVDRLLQLSQGQAVKGNQLPLVSLVRKTKTLAPNT
QTASPRALADSLMQLARQVSRLESGQ"
BASE COUNT 1009 a 1058 c 1118 g 991 t
ORIGIN
Query Match 12.0%; Score 198; DB 12; Length 4176;
Best Local Similarity 100.0%; Pred. No. 3e-32;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1456 GGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGG 1515
DB 269 GGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGG 328
QY 1516 CTGTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCGAGAAG 1575
DB 329 CTGTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCGAGAAG 388
QY 1576 CGGTCGTATAAACAGAAATTTGCTGGCGGCAGTAGCGGGTGGTCCACCTGACCCCAT 1635
DB 389 CGGTCGTATAAACAGAAATTTGCTGGCGGCAGTAGCGGGTGGTCCACCTGACCCCAT 448
QY 1636 GCCGAACCTCAGAAGTGCAA 1653
DB 449 GCCGAACCTCAGAAGTGCAA 466
RESULT 11
A29289 A29289 4239 bp DNA linear PAT 12-JUL-1995
LOCUS pSEC-Bp1 complete DNA sequence from patent DE3901681.
DEFINITION A29289
ACCESSION A29289
VERSION A29289.1 GI:1248952
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
FEATURES
Location/Qualifiers
1..4239
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 1019 a 1080 c 1137 g 1003 t
ORIGIN
Query Match 12.0%; Score 198; DB 6; Length 4239;
Best Local Similarity 100.0%; Pred. No. 3e-32;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1456 GGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGG 1515
DB 332 GGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGG 391
QY 1516 CTGTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCGAGAAG 1575
DB 392 CTGTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCGAGAAG 451
QY 1576 CGGTCGTATAAACAGAAATTTGCTGGCGGCAGTAGCGGGTGGTCCACCTGACCCCAT 1635
DB 452 CGGTCGTATAAACAGAAATTTGCTGGCGGCAGTAGCGGGTGGTCCACCTGACCCCAT 511
QY 1636 GCCGAACCTCAGAAGTGCAA 1653

```


TITLE Direct Submission
JOURNAL Submitted (22-SEP-1994) L.M. Guzman, Harvard Medical School, Dept of Microbiology & Molecular Genet., 200 Longwood Ave., Boston MA 02115, USA

REFERENCE 2 (bases 1 to 4542)
AUTHORS Guzman, L.M., Belin D., Carson, M.J. and Beckwith, J.
TITLE Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter
JOURNAL J. Bacteriol. 177 (14), 4121-4130 (1995)
MEDLINE 95332220
PUBMED 7608087

COMMENT Related sequences: J01641 and M10196 (bases 1-1300); M77749 (bases 1366-2300; 3242-4542); Gene 33, 103, 1985 (bases 2300-2784); Gene 27, 183, 1984 (bases 2784-3242).
Location/Qualifiers
1..4542
/organism="synthetic construct"
/db_xref="taxon:32630"
/lab_host="Escherichia coli"
/plasmid="PBAD24"
96..974
/gene="arac"
96..974
/gene="arac"
1003..1020
/note="Operator O2"
1125..1153
/note="PC promoter"
1161..1182
/note="Operator O1"
1204..1217
/note="CAP site"
1213..1251
/note="Operator I2 + I1"
1250..1277
/note="PBAD promoter"
1300..1366
/note="multiple cloning site"
1306..1311
/note="Optimized RBS"
1317..1319
/note="Kozak sequence"
1320..1403
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAA57431.1"
/db_xref="GI:551425"
/translation="WVPGDPLESTCRHASLAVLADERRFSA"
1367..1792
/gene="rnb"
1367..1792
/gene="rnb"
1847..1853
/gene="rnb"
1847..1853
/note="bla P3 promoter"
1885..2748
/gene="bla"
1885..2748
/gene="bla"
2784..3242
/note="M13 origin"
3248..3945
/note="pBR322 origin"
BASE COUNT 1110 a 1191 c 1136 g 1105 t
ORIGIN

Query Match 11.3%; Score 186.2; DB 12; Length 4542;
Best Local Similarity 91.6%; Pred. No. 1e-29;
Matches 197; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OY 1439 TTGACTAGTCAAGTATGGAATTGAGCTCGGTACCGGGGATCTCTAGAGTCGACCTG 1498
DB 1292 TTTTGGGCTAGCAGGAGATTCACCATGCTACCGGGGATCCTCTAGAGTCGACCTG 1351

OY 1499 CAGCATGCAAGCTTGGCTGTTTGGCGGATGAGAGAAGATTTCAGCTGATACAGATT 1558
DB 1352 CAGCATGCAAGCTTGGCTGTTTGGCGGATGAGAGAAGATTTCAGCTGATACAGATT 1411
OY 1559 AAATCAGAACGACGAGAGCGGTCTGATAAACAAGATTGCTGGCGGACGAGAGCGGCTG 1618
DB 1412 AAATCAGAACGACGAGAGCGGTCTGATAAACAAGATTGCTGGCGGACGAGAGCGGCTG 1471
OY 1619 GTCCACCTGACCCCATGCGGACCTCAGAAAGTGAA 1653
DB 1472 GTCCACCTGACCCCATGCGGACCTCAGAAAGTGAA 1506

Search completed: January 16, 2003, 08:18:44
Job time : 2967 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 06:53:16 ; Search time 253 Seconds
(without alignments)
14713.649 Million cell updates/sec

Title: US-09-830-111A-1
Perfect score: 1653
Sequence: 1 ttttggggtgaaaagtc.....atgcgaactcagaagtga 1653

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1653	100.0	1653	AAF74893	Saitoella complica
2	395.8	23.9	1353	ABN81605	Fungal decaprenyl
3	242.8	14.7	1137	AAU40138	Isoprenoid related
4	230.4	13.9	1212	AAU42942	Rhodotorula minuta
5	230.4	13.9	1278	AAU42941	Rhodotorula minuta
6	230.4	13.9	1614	AAU42940	Rhodotorula minuta
7	220.4	13.3	1134	AAU91893	Decaprenyl synthas
8	198	12.0	3753	ABA04129	Plasmid pHC19(I)
9	198	12.0	3755	ABA04130	Plasmid pHC19(II)

10	198	12.0	4176	10	AA90709	Sequence of plasmid
11	198	12.0	4189	11	AAQ05397	Secretion Vector p
12	198	12.0	4665	24	ABN84393	Antibiotic-indepen
13	195.6	11.8	1506	24	ABN81604	Fungal decaprenyl
14	187.2	11.3	1293	23	ABL03299	Drosophila melanog
15	180.8	10.9	966	21	AAQ61900	DNA encoding strai
16	180.8	10.9	966	24	ABQ62170	Arabidopsis sp. AT
17	180.8	10.9	966	24	AA596924	Arabidopsis locoph
18	179.2	10.8	5926	19	AAV32977	Tn7 donor plasmid
19	179.2	10.8	5926	19	AAV32978	Tn7 donor plasmid
20	170.4	10.3	2873	24	ABQ54840	Human ovarian anti
21	155.2	9.4	4009	15	AAQ70219	Plasmid pIG2 for T
22	152.4	9.2	3423	12	AAQ11083	Plasmid pIGRI for
23	152.4	9.2	3423	20	AAQ08615	pIGRI2 expression
24	152	9.2	1097	19	AAV39841	tac promoter/rnrb
25	152	9.2	1097	19	AAV31653	tac promoter/rnrb
26	152	9.2	1097	20	AAQ93172	p. fluorescens tac
27	148.8	9.0	6312	16	AAQ74814	Barley beta-amylase
28	148.8	9.0	6312	17	AAQ29193	Plasmid pBETA92/se
29	148.8	9.0	6312	17	AAQ03696	Plasmid pBETA92, e
30	147.8	8.9	5027	20	AAQ78874	Human tissue facto
31	147.8	8.9	5069	20	AAQ78871	Human tissue facto
32	147.8	8.9	5069	20	AAQ78896	Human tissue facto
33	147.8	8.9	5099	20	AAQ78896	Human tissue facto
34	147.8	8.9	5132	20	AAQ78883	Human tissue facto
35	147.8	8.9	5324	20	AAQ78890	Human tissue facto
36	147.8	8.9	5342	20	AAQ78893	Human tissue facto
37	147.6	8.9	6150	20	AAQ87427	Plasmid pBAD24-rec
38	147.2	8.9	518	19	AAV26002	Bacteriophage DNA
39	147.2	8.9	6806	24	ABL49925	Maltose binding prote
40	147	8.9	3474	13	AAQ29146	pMTNF-NPH plasmid
41	147	8.9	3474	14	AAQ51546	Expression plasmid
42	147	8.9	3950	22	AAQ81854	Plasmid pB8-92 nuc
43	147	8.9	3950	22	ABL54456	Nucleotide sequenc
44	147	8.9	6553	21	AAQ55456	Destination vector
45	147	8.9	8430	21	AAQ50911	Recombinant human

ALIGNMENTS

RESULT 1
AAF74893 standard; DNA; 1653 BP.

AAF74893;
23-MAY-2001 (first entry)

Saitoella complicata decaprenyl diphosphate synthase encoding DNA SEQ:1.
Saitoella complicata; decaprenyl diphosphate synthase; coenzyme Q10;

fungus; ds.
Saitoella complicata.

Key Location/Qualifiers
CDS 134..1456
/*tag= a
/product= "decaprenyl diphosphate synthase"

WO200114567-A1.

01-MAR-2001.

24-AUG-2000; 2000WO-JP05659.

24-AUG-1999; 99JP-0237561.

(KANF) KANEKA CORP.

Matsuda H, Kawamukai M, Yajima K, Ikenaka Y, Hasegawa J;
Takahashi S;

XX WPI: 2001-202937/20.
DR P-PSDB: AAB74623.
XX DNA encoding a protein having decaprenyl diphosphate synthase activity
PT and microorganism for producing coenzyme Q10 -
XX
XX
PS Claim 1; Page 23-27; 32pp; Japanese.
XX
CC The present invention describes a method for microbiologically producing
CC coenzyme Q10 at a high efficiency by using a gene of the synthesis of
CC coenzyme Q10 side chain originating in a fungus belonging to the genus
CC *Saitoella*. The present sequence encodes the specifically claimed
CC *Saitoella* complicated protein having decaprenyl diphosphate synthase
CC activity. The protein having decaprenyl diphosphate synthase activity
CC can be used for producing coenzyme Q10.
XX
SQ Sequence 1653 BP; 392 A; 417 C; 475 G; 369 T; 0 other;
Query Match 100.0%; Score 1653; DB 22; Length 1653;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTTTGTGGGGTGAAGGCGGACGGGTCAGGTTGGCTTGAGACGATAAGGCTCG 60
DB 1 TTTTGTGGGGTGAAGGCGGACGGGTCAGGTTGGCTTGAGACGATAAGGCTCG 60
QY 61 GAGATTGAGTTCAGGACCAAGCTTTGATCCGTAGGTCATCTTCAGCAAAATCATTTTC 120
DB 61 GAGATTGAGTTCAGGACCAAGCTTTGATCCGTAGGTCATCTTCAGCAAAATCATTTTC 120
QY 121 AAATCATATATACATAGGCTTCACAGCACTGGGATAGCAAGATCAGTCTTCGATCAAT 180
DB 121 AAATCATATATACATAGGCTTCACAGCACTGGGATAGCAAGATCAGTCTTCGATCAAT 180
QY 181 CGCCCTCTGGGATCGGTTTACCCCTAAGAACAGCCTCGGACCTTCATTACGACTAAGATG 240
DB 181 CGCCCTCTGGGATCGGTTTACCCCTAAGAACAGCCTCGGACCTTCATTACGACTAAGATG 240
QY 241 TACCCCGACGAGCCGCGCATCGAGTTCATGGGCTGCTGTGTCTTCGCGCTCGAGACT 300
DB 241 TACCCCGACGAGCCGCGCATCGAGTTCATGGGCTGCTGTGTCTTCGCGCTCGAGACT 300
QY 301 GGTGAGCTGTATCCGAATCAACTCTTCATCAATCCGCTCAACTTGTGTCGCGAGAT 360
DB 301 GGTGAGCTGTATCCGAATCAACTCTTCATCAATCCGCTCAACTTGTGTCGCGAGAT 360
QY 361 GTCAATCTTACATCCACATCCGATCTCTCGGTTGAGAACCCCTTCCTCGAGAC 420
DB 361 GTCAATCTTACATCCACATCCGATCTCTCGGTTGAGAACCCCTTCCTCGAGAC 420
QY 421 TGTGCTAAATCTACTATGTTTCACTGAGGAAAGCATATTCGTCGCTCATGGTACTGCT 480
DB 421 TGTGCTAAATCTACTATGTTTCACTGAGGAAAGCATATTCGTCGCTCATGGTACTGCT 480
QY 481 GATGCTCAAGCGACGAGAGTTGCGCAAAAGTTCAAGGTTGGAGAGGTGTGAGAGT 540
DB 481 GATGCTCAAGCGACGAGAGTTGCGCAAAAGTTCAAGGTTGGAGAGGTGTGAGAGT 540
QY 541 TCCGGTGAACGAGGAGCTCGACCAACAGAGTGTCAATGACAAAGAACCCAGATATGAT 600
DB 541 TCCGGTGAACGAGGAGCTCGACCAACAGAGTGTCAATGACAAAGAACCCAGATATGAT 600
QY 601 GAACATGAGGTGAGGACCTTAACGAAAGACGGGAGATCGAGGACAGACCTCGAATAT 660
DB 601 GAACATGAGGTGAGGACCTTAACGAAAGACGGGAGATCGAGGACAGACCTCGAATAT 660
QY 661 CCTGCTTCGCAACGCGGTTGGCTGAGATCAAGAGATGATTCATGCACTACTCT 720
DB 661 CCTGCTTCGCAACGCGGTTGGCTGAGATCAAGAGATGATTCATGCACTACTCT 720
QY 721 CCACGACGACTTATTCAGGCTTCGAGACGAGCAAAACGACCATCCGAAACACAGGC 780
DB 721 CCACGACGACTTATTCAGGCTTCGAGACGAGCAAAACGACCATCCGAAACACAGGC 780

DB 721 CCACGACGACTTATTCAGGCTTCGAGACGAGCAAAACGACCATCCGAAACACAGGC 780
QY 781 ATTGGGAAACAGATGCGGATTTTGGCTGGTGAATTTCTTGTGGAGCGGCTGTGTC 840
DB 781 ATTGGGAAACAGATGCGGATTTTGGCTGGTGAATTTCTTGTGGAGCGGCTGTGTC 840
QY 841 APTGGCGAGTGGGCAATCCGGAGGATGATGACTTTTGGCTGCTGATATGCAAACTT 900
DB 841 APTGGCGAGTGGGCAATCCGGAGGATGATGACTTTTGGCTGCTGATATGCAAACTT 900
QY 901 GGTGAGGAGAGTTCATGCACTTGAATAAATATCTTGATGATCCGATGAGCTACGCG 960
DB 901 GGTGAGGAGAGTTCATGCACTTGAATAAATATCTTGATGATCCGATGAGCTACGCG 960
QY 961 GACGAGGAAACGTTGATCTATTTTGGAGAGACTTACTTGAAGACTGCTCTGAT 1020
DB 961 GACGAGGAAACGTTGATCTATTTTGGAGAGACTTACTTGAAGACTGCTCTGAT 1020
QY 1021 TGGCAAGTCGACAGAGAACTGGGCTTGGGAGTGAAGCTACGCTGAGTGGCTGATGC 1080
DB 1021 TGGCAAGTCGACAGAGAACTGGGCTTGGGAGTGAAGCTACGCTGAGTGGCTGATGC 1080
QY 1081 TGTATGCTTACGGAAGAACTTGGTTTGGCATTCAGATCGTCGACGACATGCTCGA 1140
DB 1081 TGTATGCTTACGGAAGAACTTGGTTTGGCATTCAGATCGTCGACGACATGCTCGA 1140
QY 1141 CTACACGCTCCGCTACGACCTCGGTAAGCCGCGGTGACAGCTCCAGCTCGGCT 1200
DB 1141 CTACACGCTCCGCTACGACCTCGGTAAGCCGCGGTGACAGCTCCAGCTCGGCT 1200
QY 1201 CGCCACGCGCGCGGCTCTTGGATGAGAACACACGCGGAGCTCGCTCCATGATCAA 1260
DB 1201 CGCCACGCGCGCGGCTCTTGGATGAGAACACACGCGGAGCTCGCTCCATGATCAA 1260
QY 1261 GCGCAAGTCTCTGACCCGAGAGAGCTGAGCGCTGCAAGCGAGTTGTGAGAAAGTGA 1320
DB 1261 GCGCAAGTCTCTGACCCGAGAGAGCTGAGCGCTGCAAGCGAGTTGTGAGAAAGTGA 1320
QY 1321 TGGATTGAGAAACGAGAGCTTGGCGGAGAGATGCCCCAGAGGCTTGATGCAAT 1380
DB 1321 TGGATTGAGAAACGAGAGCTTGGCGGAGAGATGCCCCAGAGGCTTGATGCAAT 1380
QY 1381 TGGAGCTTCCCGAGAGCTCGGACCGAAGGCTTGGAGAGAGTGAACGACAGGCTT 1440
DB 1381 TGGAGCTTCCCGAGAGCTCGGACCGAAGGCTTGGAGAGAGTGAACGACAGGCTT 1440
QY 1441 GACTAGTCAAGATAGGAATTCGAGCTCGATACCGGGGATCTCTAGAGTCACTGCA 1500
DB 1441 GACTAGTCAAGATAGGAATTCGAGCTCGATACCGGGGATCTCTAGAGTCACTGCA 1500
QY 1501 GGCATGCAAGCTTGGCTGTTTGGCGGATGAGAAAGATTTTCAGCTGATACAGATTA 1560
DB 1501 GGCATGCAAGCTTGGCTGTTTGGCGGATGAGAAAGATTTTCAGCTGATACAGATTA 1560
QY 1561 ATCGAAGCGAGAGCGGTCTGATTAACAGATTTTGGCTGCGGAGTACCGGGTGT 1620
DB 1561 ATCGAAGCGAGAGCGGTCTGATTAACAGATTTTGGCTGCGGAGTACCGGGTGT 1620
QY 1621 CCCACCTGACCCCATGCGCACTCAGAAAGTGA 1653
DB 1621 CCCACCTGACCCCATGCGCACTCAGAAAGTGA 1653
RESULT 2
ID ABN81605
XX ABN81605 standard; DNA; 1353 BP.
AC ABN81605;
XX
XX 29-AUG-2002 (first entry)
XX
XX Fungal decaprenyl diphosphate synthase encoding DNA SEQ ID NO 2.
XX

Qy	387	CTCTCCTCGGTTGAGGACACCCCTTCTCTCGACACTGTCGTTAAATACTATGTTCACTGCTG 446
Db	65	CCCTACTTGGGATCAGGTCATCCCGCCCTAGACACGATAGCAAAGTACTACTTCCAAGCGG 124
Qy	447	AGGGAAGCATATTCTGTCCTCATGGTACTGCTGATGGCTCAGGCGACGA-GGTTGCG 505
Db	125	AGGCAAAACATATTGGCCTATGATGTTCTCATGTGCCAACCAACCGGTCTAG 184
Qy	506	CCAAAGTTGAGGGTTGGGAGAAGTCTGTGGAGGTTTCGGTGAACGAGGGACTCGCACCA 565
Db	185	CGCCCGGTTTGAAGAACGCTCAAAATTGGAACTATCAGGTCGGAACAGACTGATCCCT 244
Qy	566	CCAG-----AGGTGCTCAATGACAAGAACCCAGATATGATGAACATGAGGTCAGAC 617
Db	245	CCAGTCAATCATGATCTCTCGAAGTGAAGCAGATGAGATACTCAACGATTTCGAATC 304
Qy	618	CATTAAACGAAGGACG-----GGCAGATCGAGGACAGACGTCGAATA 659
Db	305	CCTCTCTGTTGCGTGGAGCTCCTCTTCGCGCTCGATAGCATGCGGTCCAGTCGAAATG 366
Qy	660	TCCTCGCCTCGAAACGCGGTGGCTGAGATCAACGGAGATGATCATCGAGGATCACTCC 719
Db	365	TCCTACCTCGCAACGACGCTCGCGAAATCACCGAAATGATCAACGATTTCGAT 424
Qy	720	TCCACGACGAGTTTATCGACGCTTCGAGACGAGAAACGACCATCCGGAACCCAGG 779
Db	425	TGCACGACGATGTCATAGACGTTTCAGCATGAGGAGACCAACGCTCCGCCCCCGCTG 484
Qy	780	CATTTCGAAACAAAGATGGCGATTTCGGCTGGTGATTTCTTGTGGAGCGGCGTCTGTTG 839
Db	485	CATTTCGGGAACAAGTCTCGTCTCGCGGGGATTTCTCTCTCGCTCGTCTCGCTGT 544
Qy	840	CATTGGCGAGTTTCGCAATTCGCGAGGTGATTGAGCTTTTGGCTACTGTTATGCAAACT 899
Db	545	ACCTCTCCGACTAGGAGCAACGAGGTGCTCGAGTAGTACCTCCGTGTAGCTAATC 604
Qy	900	TGTTTGAGGAGAGTTTCATCGAGTCTGAAATACTGTTGATGATCGGATT----- 949
Db	605	TAGTAGAGGGAAGTCATCGAGATCAAGGGAATGCTCTTGAAGCAATGCAAGCGGAA 664
Qy	950	-----GAGGCTACGGCGACCGAGAAACGTTTCGATTACTATTTGAGAGA 995
Db	665	GCAAAGAGTAGCAGTGCACAGATTGACCCCGAAATTTTCGAACTATATGAAGA 724
Qy	996	CTTACTTTGAAGCTCGTCTTGAATGCCAGTCTGTGACAGCAAGTCGGCTCTCGGGT 1055
Db	725	CATACTTGAAGACCGCAAGTCTCATCGGAAATTCGAACAGAGCGACCACTATCTCGGTG 784
Qy	1056	GTGCTACGCTGA-----GGTTGCTCATGCTCGTTATGCTT 1091
Db	785	GAGCAGCGGAGAACAGGGGTGGATAGAGGCGAGGCGCAATAAGACATTCGGTACTCGT 844
Qy	1092	ACGAAGGAACCTTGGTTTGGCAATCCAGATCGTCGACGACATGCTCGACTACACCGTCT 1151
Db	845	ACGGTCGCAATCTAGGTATTGCTTTCCAGCTCGTCGACGATCTACTAGATTTACAGCTA 904
Qy	1152	CCGCTACCGACCTCGGTAAACCCGC---CGGTGACAGCTTCGAGTCTCGGTCTCGCACCG 1208
Db	905	CAGACGCGCAATTCGGCAAGCCCTCACAGGTCAGATCTGAAGTCTGGTCTCGCAACTG 964
Qy	1209	CGCCGGCCCTTTCGATGAAGCACACCGCGAGCTCGGTCCCATGATCAAGCGCAAGT 1268
Db	965	CGCCCGGCTGTACGATGGAAGAGGTTCCCGAGATGGGCCAGATGATTCCTCGCAAGT 1024
Qy	1269	TCTCTGACCCAGGAGCGTCGACGCTGACGCGAGTTGTCGAGAAAAAGTGATGGATTGG 1328
Db	1025	TTGAAGACGAAGGCGATGTCGAACTGCCAGGAACTCTAGTAAGAAAGTCAGCTGACCGG 1084
Qy	1329	AGAAGAAGAGCCTTGGCGGAGGATGTCGCCAGNAGCGGTTTGATGCAATTCGAGAGT 1388
Db	1085	AAAAAGACCGTGAATTTGGCGGAAAAATTCGCGCACTCGCAATGAGGCCCTCGAGGAT 1144

Qy	1399	TCCCGGAGATCCGGACCGAAGGCTTTGGAGCAGTGTGACGACAAAGTGTGTGACTAGGT	1449
Db	1145	TGCCGGAGTCCGACGCTAGAGAAAGCGCTCGAAGGCGCTGACCAAGACTGTGCTCAACCGAA	1204
Qy	1449	CAAGATAG	1456
Db	1205	CAAGTAG	1212
RESULT 5			
ID	AAL42941	standard; DNA; 1278 BP.	
XX	AAL42941;		
XX	08-AUG-2002	(first entry)	
XX	Rhodotorula minuta decaprenyl diphosphate synthase 2 coding sequence.		
XX	Fungi; decaprenyl diphosphate synthase; gene; ds; enzyme;		
KW	high-yield fermentation; coenzyme Q10 production.		
XX	Rhodotorula minuta.		
XX	Key	Location/Qualifiers	
FH	CDS	1..1278	
FT		/tag= a	
FT		/product= "Rhodotorula minuta decaprenyl diphosphate	
FT		synthase 2"	
XX	WO200240682-A1.		
XX	23-MAY-2002.		
XX	20-NOV-2001; 2001WO-JP10119.		
XX	20-NOV-2000; 2000JP-0352940.		
XX	(KANF) KANEKA CORP.		
XX	Matsuda H, Kawamukai M, Yajima K, Ikenaka Y;		
XX	WPI; 2002-427097/45.		
DR	P-PSDB; AAO14998.		
XX	Decaprenyl diphosphate synthase gene of Rhodotorula origin for		
PT	efficient preparation of coenzyme Q10		
XX	Claim 2; Page 36-39; 50pp; Japanese.		
XX	The invention comprises the amino acid and coding sequences of fungal		
CC	(Rhodotorula minuta) decaprenyl diphosphate synthase enzymes. The fungal		
CC	decaprenyl diphosphate synthase DNA and protein sequences are useful for		
CC	the efficient, high-yield fermentative production of coenzyme Q10 - for		
CC	pharmaceutical use. The present DNA sequence encodes a Rhodotorula min		
CC	decaprenyl diphosphate synthase enzyme.		
XX	Sequence 1278 BP; 345 A; 349 C; 334 G; 250 T; 0 other;		
SQ	Query Match	13.9%; Score 230.4; DB 24; Length 1278;	
	Best Local Similarity	54.1%; Pred. No. 3e-49;	
	Matches	654; Conservative 0; Mismatches 476; Indels 78; Gaps	
Qy	327	TCATCAATCCGCTCAACTGGTCCGTCGGAGATGTCAAATCTTACATCCACATCCGAT	386
Db	71	TCCTCGACCCACCTCCAACTCGTAGGCAATGAATGTCAAGTCTCCGGTCAACCGTCCAAG	130
Qy	387	CTCTCCTCGGTTCAGGACACCCCTTCTCGACACTGTCGCTAAATCTATGTTGAGTCTG	446
Db	131	CCCTACTTGGGATCAGGTATCCCGCCCTAGACACGATAGCAAGTACTTCTTCAAGGGG	190
Qy	447	AGGGAAGACATATTCTGTCGCTCATGCTACTGCTGATGGCTCAGGCGACGGA-GGTTGCG	505

```

Db 191 AGGGAACATATTTGGGCTATGATCGTTCTTCATGTCACCAAGCAAAAGGCTAG 250
QY 506 CCAAAAGTTCAAGGTTGGGAGAAAGTCGTGAGGTTCCGGTGAACGAGGACTCCGACCA 565
Db 251 CGCCGGGGTTTGAAGAACGCTCAAAATTGGAATCATGAGTGGGAAAACAGACTGATCCCT 310
QY 566 CCAG-----AGTGCTCAATGACACAGAACCCAGATATGATGAAACATGAGTCAAGAC 617
Db 311 CCAAGTCAATCAATATCTCTCGAAGTGAAGACAGATGAGTACTCAACGATCGAATC 370
QY 618 CATTAACGAGAGCG-----GCGAGATCCAGGAGCAAGAGCTCGAATA 659
Db 371 CCTCTTCGTTGCTGCTGAGAGCTCCTCTTGCCGCTGATAGCATGCCCTCCAGCTCGAATG 430
QY 660 TCCCTGCTCGCAACGCGCGGTGGCTGAGATCAAGAGATGATCCATGACAGCATCTCC 719
Db 431 TCCTAACCTCGCAACGAGCCTCGCGGAAATCACCGAAATGATCCAGTACGCTTCTAT 490
QY 720 TCCACGACGACGTTATCCAGCGCTTCGAGACCAAGCAAAACGACCATCCGGAACCAAG 779
Db 491 TGCACGACGATGTCATGACGCTTCAGCCATGAGAGACACAAAGGTCGCCGCCGCTG 550
QY 780 CATTCGGAACAAAGATGGGATTTGGCTGATGATTTCTTGTGGGACGGGCTGTTG 839
Db 551 CATTCGGGAACAAGATCTGCTGCTGCGGCGGATTTCTCTCTGCTGCTGCTGCTGCT 610
QY 840 CATTCGCGAGTTGGCGCAATCCGAGGTGATGAGCTTTGGCTAATGTTATGCAACT 899
Db 611 ACCTCCCGGACTAGGAGCAACGAGGTGCTGAGCTAGTACCTCCGCTAGCTAATC 670
QY 900 TGGTTGAGGAGAGTTTCATGACGCTGAAAAATATCTGTTATGATGCGANT----- 949
Db 671 TAGTAGAGGCGCAAGTCATGACATCAAGAAATGCTCTGAAAGCAATGCAAGCGGAA 730
QY 950 -----GAGGCTAGCGGAGCGGAGAAACGTTCAATTAATTTTGCAGAAAG 995
Db 731 GCAAAAGGTAGCAGTGACAGATTCGACATTCGACGAAATTTTCGAACTTATGAAAGAA 790
QY 996 CTTACTTGAAGACTGCGTCTTGTATTCGCAAGTCTGACAGCAAGTGGCTTCTGGGTG 1055
Db 791 CATACTTGAAGACCGCAATCTCATTCGCAAAATCGCAAGAGGACCACTATCTCGGTG 850
QY 1056 GTGTAACGCTGA-----GGTTGCTATGCTGCTTATGCTT 1091
Db 851 GAGCAGGCGAGAAACAGGGGTGATAGAGGGCGAGCGCATTAAGACATTTGGGTCTCGT 910
QY 1092 ACGGAAGAACTTGTGTTGGCATTCAGATCGTGGAGCAATGCTCGACTACCGGTCT 1151
Db 911 ACGGTGCAATGTAGGTATGCTTTCACACTGCTGACGATCTACTAGATTTTCACAGCTA 970
QY 1152 CGGCTACGAGCTCGGTGTAAGCCCGC---CGGTGACAGCTCCAGCTCGGCTCGCCACCG 1208
Db 971 CAGAGCGCGAATTCGCGCAAGCCCTCACAGGTGCAATCTGAAGCTCGGTTCTCGCAACTG 1030
QY 1209 CGCGCGGCTCTTTCGATGAAAGCACACGCGGAGCTCGGTCCCATGATCAAGCGCAAT 1268
Db 1031 CGCCGCGCTGTACGATGGAAGATTCGCGAGATGGGCGAGATGATTTCTCCGCAAGT 1090
QY 1269 TCTTCGACCCAGAGACCTGACGCTGACGCGAGTGGTTCAGAAAAATGATGATGTTGG 1328
Db 1091 TTGAGAACCAAGCGCGATGTCGAAATGCGCAGAAATCTAATAAGAAAGTCAAGCTGGACCG 1150
QY 1329 AGAAGACGAGAGCTTTGGCGGAGAGATATGCCAGAAAGCGTTGGATTCGAGACT 1388
Db 1151 AAAAAGCCTGAATTTGGCGGAAAAACATGCCGCACTCGCATGAGGCGCTCGAGGAT 1210
QY 1389 TCCCGGAGAGTCCGCGACCGAAGGCTTTGAGAGATGTAACGCAAGGTTGTAAGGT 1448
Db 1211 TGCCGAGAGTCCGAGCGCTAGAGAAAGCGCTGAAAGGCTGACCAAGACTGTGCTCAACGAA 1270
QY 1449 CAAGATAG 1456
Db 1271 CAAGATAG 1278

```

```

RESULT 6
AAL42940
ID AAL42940 standard; DNA; 1614 BP.
XX
XX
AC AAL42940;
XX
XX 08-AUG-2002 (first entry)
XX
XX Rhodotorula minuta decaprenyl diphosphate synthase 1 coding sequence.
XX
XX Fungi: decaprenyl diphosphate synthase; gene; ds; enzyme;
XX high-yield fermentation; coenzyme Q10 production.
XX Rhodotorula minuta.
XX
XX Key Location/Qualifiers
XX CDS 1..1614
XX FT /*tag= a
XX FT /product= "Rhodotorula minuta decaprenyl diphosphate
XX FT synthase 1"
XX
XX MO200240682-A1.
XX
XX PD 23-MAY-2002.
XX
XX PF 20-NOV-2001; 2001WO-JP10119.
XX
XX PR 20-NOV-2000; 2000JP-0352940.
XX
XX PA (KANF ) KANEKA CORP.
XX
XX PI Matsuda H, Kawamukai M, Yajima K, Ikenaka Y;
XX WPI: 2002-427097/45.
XX DR P-PSDB; AMO14997.
XX
XX PT Decaprenyl diphosphate synthase gene of Rhodotorula origin for
XX efficient preparation of coenzyme Q10
XX
XX PS Claim 1; Page 29-33; 50pp; Japanese.
XX
XX CC The invention comprises the amino acid and coding sequences of fungal
XX (Rhodotorula minuta) decaprenyl diphosphate synthase enzymes. The fungal
XX decaprenyl diphosphate synthase DNA and protein sequences are useful for
XX the efficient, high-yield fermentative production of coenzyme Q10 - for
XX pharmaceutical use. The present DNA sequence encodes a Rhodotorula minuta
XX decaprenyl diphosphate synthase enzyme.
XX
XX SQ Sequence 1614 BP; 432 A; 462 C; 377 G; 343 T; 0 other;
XX
XX
XX Query Match 13.9%; Score 230.4; DB 24; Length 1614;
XX Best Local Similarity 54.1%; Pred. No. 3.3e-48;
XX Matches 654; Conservative 0; Mismatches 476; Indels 78; Gaps 6;
XX
QY 327 TCATCAATCCGCTCAACTTGTGCGTCCCGAGATGTCAAATCTTAATCAATCCAACTCCGAT 386
Db 407 TCTTGACCACTCCAACTGTAAGCAATGAAGTCAAGTCTCCGTCAAACGTCGAAG 466
QY 387 CTCTCTCGGTTCCAGACACCTTCTCTGACACTGTCGTAATTAATTAATTAATTAATG 446
Db 467 CCTTCTGAGGATCAAGTATCCGCTCCGCTGACAGATGCAAGATCAATTAATTAATTAAT 526
QY 447 AGGGAACGATATTCGTCGCTGATGATGATGATGATGATGATGATGATGATGATGATG 505
Db 527 AGGGGAACATATTTGGGCTTATGATGTTCTTCATATCCCAAGCCCAACGATCTAG 586
QY 506 CCAAAAGTTCAAGGTTGGGAGAAAGTCCGTGAGGTTCCGTTGAACGAGGACTCCGACCA 565
Db 587 CGCCGGGGTTTGAAGAACGCTCAAAATTTGAACATATCAGGTGCGAAACAGACTGATCCCT 646
QY 566 CCAG-----AGTGCTCAATGACAAAGAACCCAGATATGATGAAACATGAGTCAAGAC 617

```

```

Db 647  ||||  CCAGGTCAATCAATGATCTCTCGAAGTGAAGCAGATGAGATCAACGATTCGAATC 706
QY 618  ||||  CATTACGAGGAGC-----CGAGATCGAGGACAGACGTCGGAATA 659
Db 707  ||||  CCTCTTCGCTCGCTCGAGCTCTCTTCGCGCTCGATAGCATCCGCTCCACGTCGAATG 766
QY 660  ||||  TCCTCGCTCGCAACGCGGTTCGCTGAGATCAGGAGATGATCCATCGACATCATCC 719
Db 767  ||||  TCCTACCTCGCAACGAGCCTCGCGGAATCACCGAAATGATCCAGTAGCTTCGCTAT 826
QY 720  ||||  TCACACGACGCTTATCGAGCTTCGAGACCAAGCAACGACCACTCCGGAACAGG 779
Db 827  ||||  TGCACGAGATGTCATAGACGTTACCCATGAGGAGACACAGCTCGCCGCCCTG 886
QY 780  ||||  CATTGGAACACAGATCGGATTTTGGCTGGTGAATTTCTTTGGGACGGGCGTCTGTTG 839
Db 887  ||||  CATTGCGGAACAGATCTCGTGTCTGGCGGGATTTCTCTCGCTCGTCTCGCTGT 946
QY 840  ||||  CATTGGCGAGTTGCGCAATCCGAGGTGATTGAGCTTTTGGCTACTGTTATTGCAAACT 899
Db 947  ||||  ACCTCTCCGACTAGGAGCAACGAGCTGTCGAGCTAGTAGCTCGCTGCTAGCTAATC 1006
QY 900  ||||  TGCTGAGGAGAGTTCATGCACTGAAATAACTGTTGATGATGCGATT----- 949
Db 1007  ||||  TAGTAGAGGCGAAGTCATGCAAGTCAAGGAATGCTCTGAAAGCAATGCAAGCGGAA 1066
QY 950  ||||  -----GAGGTACCGCGACGCGAGGAACGTTGCTGATTACTATTTCGAGAAGA 995
Db 1067  ||||  GCAAAGAGTAGCAGTGCACAGATTGACCCCGGAAATTTTCGAACTATTATGAGAAGA 1126
QY 996  ||||  CTACTTGAAGACTGCTCTTGTATTCGAAAGTCTGCGAGCAAGTGCCTTCGGGTG 1055
Db 1127  ||||  CATACTTGAAGCGCAAGTCTCATCGGAAATCGACAAGAGCGACCACTATCTCGGTG 1186
QY 1056  ||||  GTGCTACGCTGA-----GGTGTCTGATGCTCTTATGCTT 1091
Db 1187  ||||  GAGCAGCGGAACACAGGGTGGATAGAGGCGAGCGCATAAAGACATTCGCTACTCGT 1246
QY 1092  ||||  ACGGAAGAACCTTTGGCTATTCAGATCGTCGACGACATGCTCGACTACACCTCT 1151
Db 1247  ||||  ACGTGCATCTAGGTATTGCTTTCCAGCTCGTCGACGATCTACTAGATTACAGCTA 1306
QY 1152  ||||  CCGTACCGACCTCGTTAAGCCGC-----CGTGCAACCTCCAGCTCGGTCTGCCACCG 1208
Db 1307  ||||  CAGACGCGCAATTGCGCAAGCCTTCACAGGCTGAGATCTGAAGCTCGGTCTCGCAACTG 1366
QY 1209  ||||  CGCGGCGCTCTTCGCATGGAAGCACACCGAGCTCGGTCCCATGATCAAGCGCAAGT 1268
Db 1367  ||||  CGCGCGCTGTACGATGGAGAGATTCCCGGAGATGGGCGGAGATTTCTCGCAAGT 1426
QY 1269  ||||  TCTCTGACCCAGGAGAGCTCGAGCGTGCACGCGAGTTGGTTCGAGAAAAGTGATGGTGG 1328
Db 1427  ||||  TTGAGACGAAGCGATGTCGAATCTCCAGGAATCTAGTAAGAAAGTCACTGGACCGG 1486
QY 1329  ||||  AGAAGACGAGAGCCTTCGCGGAGAGATGATCCAGAGGCGTTGGATGCAATTGCGACGT 1388
Db 1487  ||||  AAAAGACCGTGAATTTGGCGGAAAAAATGCGCACTCGCAATGAGGCGCTCGAGGAT 1546
QY 1389  ||||  TCCGAGAGTCCGCGACGGAAGCTTTGGAGCAGTTGACGACAGGTGTTGACTAGGT 1448
Db 1547  ||||  TCGCGAGTCCGAGCTGAGAGAGCGCTCGAAGCGCTGCAACGACTGTGCTCAACCGAA 1606
QY 1449  ||||  CAAGATAG 1456
Db 1607  ||||  CAAGTAG 1614

```

RESULT 7
 AAT91893
 ID AAT91893 standard; DNA; 1134 BP.
 XX
 AC AAT91893;

```

XX 12-MAR-1998 (first entry)
XX Decaprenyl synthase gene.
XX Decaprenyl synthase gene; heterogenic ubiquinone production;
KW transferase; mitochondrial; yeast; hexaprenyl diphosphate;
KW modification; genetic engineering; ds.
XX Schizosaccharomyces pombe.
XX Key Location/Qualifiers
FT CDS 1..1134
FT /tag= a
FT /product= Decaprenyl_synthase
XX JP09173076-A.
XX 08-JUL-1997.
XX 27-DEC-1995; 95JP-0351243.
XX 27-DEC-1995; 95JP-0351243.
XX (ALPH-) ALPHA SHOKUHN KK.
XX WPI; 1997-397034/37.
XX P-PSDB; AAW30762.
XX New method for heterogenic ubiquinone production - allows the side
XX chain of ubiquinone to be freely modified by genetic engineering
XX methods
XX Claim 9; Page 6; 8pp; Japanese.
XX A new method has been developed for the formation of heterogenic prenyl
XX transferase. The method involves the fusing of a fragment containing a
XX hexaprenyl diphosphate synthase structural gene Coq1 to a heterogenic
XX prenyl transferase gene and expressing the resultant fused gene. The
XX present sequence encodes a specifically claimed decaprenyl synthase.
XX The method allows the side chain of ubiquinone to be freely modified
XX by genetic engineering methods.
XX Sequence 1134 BP; 328 A; 220 C; 244 G; 342 T; 0 other;
Query Match 13.3%; Score 220.4; DB 18; Length 1134;
Best Local Similarity 56.5%; Pred. No. 1.1e-46;
Matches 451; Conservative 0; Mismatches 341; Indels 6; Gaps 2;
QY 649 GAGCTCGAATATCTCGCTCGCAACGCGGTTGGCTGAGATCAGGAGATGATCCATGC 708
Db 333 GACGGTCAAAATCTTCCCTCTCAATTGAGATTAGACAATAACCGAGATGATCCATAT 392
QY 709 AGATCACTCTCCACGACGCTTATCGACGCTTCGAGACCAGACGAAACCGCACATC 768
Db 393 AGCAAGTTTGCTGCATGACGATGTGATGATCAGCTTAATGTCGTAGAGGCTCACCTTC 452
QY 769 CGAAACACGAGCATTCGGAACCAAGATGGGATTTTGGCTGGTGAATTTCTTTGGGACG 828
Db 453 AAGCAATGTGTCTTCGGTAATCGACCGTCAATCTTCGGGGTAATTTCTCTTCGACG 512
QY 829 GGGCTGTGTTCGATTCGCGAGGTTGCGCAATCCGAGGTCATTGAGCTTTTGGCTACTGT 888
Db 513 GGCTTCGACTGTATGGCCGCTTCGAAATCCCAAGTTACGAGTTGTTAGCTACAGT 572
QY 889 TATTGCAAACTTGGTTGAGGAGAGATTTCATGCGAGTTGAAAAAATCTGTTGATGATCGGAT 948
Db 573 GATAGCAGACTTGGTTTCGAGGTGAGTTTTCGAGCTAAAAAATCTATGATCCTTCATC 632
QY 949 TGAGGCTACGCGACGAGGAAACGTTTCGATTACTATTTCAGAGACTTACTTGAAGAC 1008
Db 633 TTTGGAAA---TAAAAACAATCAAAATTTTGACTATTATTGAAAAAAGTTTGTGAA--- 686

```


QY 1009 TCGCTCTTGATTCGCAAGCTGTCAGAGCAAGTCCGCTTCTGGGTGGTGTCTACCGCTCA 1068
DB 687 ACAGCGCACTTATTTCCAAAGCTGCAAGGCTTCTCAATCCCTCGACACAAATGTTCTCTCA 746
QY 1069 GGTTCCTCATGTGCTGTTATGCTTACGAAAGAACCTGGTGGATTCCAGATCGTCGA 1128
DB 747 TGTAGCAACAGCTGCTGGGGAATACGCTGATGATGCAATGCTTCTTCAACTAATGGA 806
QY 1129 CGACATGCTCGACTACACCGCTCCGCTACCGACCTCGGTAAACCCCGCGATGACACCT 1188
DB 807 TGACCTGTTGACCTAATGCTGCAAAAGATGATCTTAGGAAAGCGGCTGGTGACGATTT 866
QY 1189 CCAGCTCGCTTCGCAACCGCCGCGCTCTTCGATGAAAGCACACCGCGAGCTCG 1248
DB 867 GAAGCTAGGGTGGCTACAGCTCCGCTCTCTTGATGAAAGATATCCAGAACTTGG 926
QY 1249 TCCCATGATCAAGCGCAAGTCTGACCCAGAGACGCTGACCGTCAACGCGAGTTGGT 1308
DB 927 TGCAATGATGTAATGAAATGATCAATCATCTTCTGATATCCAAAGGCTGGTCTTGGT 986
QY 1309 CGAGAAAGTGAATGATGAGAGAGACGAGCTTGGCGAGAGATATGCCAGAAAGC 1368
DB 987 TGAGTGCACATGATGCTATGACGACAAACATCACTTGGGCAAAAGATATCAAAAAAGC 1046
QY 1369 GTTGATGCAATTCGAGCGTTCGCGAGAGTCCGCAACGGAAGCTTTGGAGCAGTTGAC 1428
DB 1047 CAAAGATTCCTTCTGCTGCTCCCTGATTCACCTGCAAGAGAGACCTTTTGGCGTTGC 1106
QY 1429 GACAAAGTGTGACTAG 1446
DB 1107 TGATTAAGTAATACGAG 1124

RESULT 8

ABA04129
ID ABA04129 standard; DNA; 3753 BP.

AC ABA04129;

DT 28-FEB-2002 (first entry)

DE Plasmid pHCE19T(II) polynucleotide sequence SEQ ID NO:3.

XX Promoter; gene expression; fermentation; D-AAT; Bacillus SK-1;

KW D-amino acid aminotransferase; economic; large-scale protein production;

KW circular; ds.

OS Bacillus sp. SK-1.

OS Synthetic.

XX Key

FT promoter

FT /tag= a

FT /note= "specifically claimed promoter region in

FT Claim 1 (see ABA04127)"

XX WO200183787-A1.

XX 08-NOV-2001.

XX 26-APR-2001; 2001WO-JP03607.

XX *27-APR-2000; 2000JP-0128528.

XX (BIOL-) BIOLEADERS CORP.

XX (TAKI) TAKARA SHUZO CO LTD.

XX Sung M, Lee S, Hong S, Seo H;

XX WPI; 2002-066535/09.

XX Promoter sequence effective in Escherichia coli and Bacillus for

PT economic large-scale fermentative production of proteins
XX Claim 7; Page 55-57; 65pp; Japanese.
PS The present invention describes a DNA sequence (I) comprising promoter
XX activity in Escherichia coli or Bacillus cells, where the promoter is
CC derived from the D-AAT (D-amino acid aminotransferase) gene of
CC Bacillus SK-1. (I) can be used for the high level expression of a
CC foreign gene in a bacterial host for economic and efficient large-scale
CC production of proteins such as enzymes, cytokines and antibodies. The
CC present sequence represents the plasmid pHCE19T(II) polynucleotide
CC sequence comprising a specifically claimed promoter region from the
CC present invention.
XX Sequence 3753 BP; 953 A; 912 C; 929 G; 959 T; 0 other;
SQ

Query Match 12.0%; Score 198; DB 24; Length 3753;

Best Local Similarity 100.0%; Pred. No. 9.4e-41;

Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1456 GGAATTCGAGCTCGGTACCCGGGATCCTTAGAGTCGACCTGCAAGCATCAAGCTTGG 1515

DB 226 GGAATTCGAGCTCGGTACCCGGGATCCTTAGAGTCGACCTGCAAGCATCAAGCTTGG 285

QY 1516 CTGTTTGGCGGATGAGAGATTTTCAGGCTGATACGATTAATCAGAACGACAGAG 1575

DB 286 CTGTTTGGCGGATGAGAGATTTTCAGGCTGATACGATTAATCAGAACGACAGAG 345

QY 1576 CGGTCTGATTAACAGAAATTTGCTGGCGGAGTAGCGGCTGCTCCCACTGACCCCAT 1635

DB 346 CGGTCTGATTAACAGAAATTTGCTGGCGGAGTAGCGGCTGCTCCCACTGACCCCAT 405

QY 1636 GCCGAATCTCAGAGTGAA 1653

DB 406 GCCGAATCTCAGAGTGAA 423

RESULT 9

ABA04130
ID ABA04130 standard; DNA; 3755 BP.

AC ABA04130;

DT 28-FEB-2002 (first entry)

DE Plasmid pHCE19(II) polynucleotide sequence SEQ ID NO:4.

XX Promoter; gene expression; fermentation; D-AAT; Bacillus SK-1;

KW D-amino acid aminotransferase; economic; large-scale protein production;

KW circular; ds.

OS Bacillus sp. SK-1.

OS Synthetic.

XX Key

FT promoter

FT /tag= a

FT /note= "specifically claimed promoter region in

FT Claim 1 (see ABA04128)"

XX WO200183787-A1.

XX 08-NOV-2001.

XX 26-APR-2001; 2001WO-JP03607.

XX *27-APR-2000; 2000JP-0128528.

XX (BIOL-) BIOLEADERS CORP.

XX (TAKI) TAKARA SHUZO CO LTD.

XX Sung M, Lee S, Hong S, Seo H;

XX WPI; 2002-066535/09.

XX Promoter sequence effective in Escherichia coli and Bacillus for

XX 21-JAN-1989; 89DE-3901681.
 XX (BEHW) BEHRINGWERKE AG.
 XX
 XX Knapp S, Abel KJ;
 XX WPI; 1990-232260/31.
 DR P-PSDB; AAR96422.
 XX
 XX Signal peptide from Bordetella pertussis - causing secretion of
 PT heterologous proteins in E.coli, and expression vectors for isolating
 PT and testing signal sequences.
 XX
 XX Disclosure; ; 8pp; German.
 XX
 XX Complete sequence of pSEC-BP1. Constructed from a 3.1kb SacI-ScaI
 CC fragment of pTRC99C-phoA-seq1 (including signal sequence) ligated to
 CC a 0.9kb SacI-ScaI fragment from pTRC97A. DNA encoding an
 CC heterologous protein can be inserted into the polylinker region.
 CC Upstream of this is a strong promoter (perfect match to consensus
 CC -35 and -10 regions) and the B.pertussis signal sequence.
 CC Downstream are three stop codons, one in each possible reading
 CC frame.
 CC See also AA005398-Q05400 and AA005521-Q05526.
 CC
 XX Sequence 4189 BP; 1004 A; 1067 C; 1123 G; 995 T; 0 other;
 SQ
 Query Match 12.0%; Score 198; DB 11; Length 4189;
 Best Local Similarity 100.0%; Pred. No. 9.8e-41;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1456 GGAATTTCGAGCTCGGTACCCGGGGATCCTTAGAGTCGACCTGCAGCGCATGCAAGCTTGG 1515
 DB 332 GGAATTTCGAGCTCGGTACCCGGGGATCCTTAGAGTCGACCTGCAGCGCATGCAAGCTTGG 391
 QY 1516 CTGTTTGGCGGATGAGAGAAATTTTCAGCTGATACGATTAATTCGAACGCAAG 1575
 DB 392 CTGTTTGGCGGATGAGAGAAATTTTCAGCTGATACGATTAATTCGAACGCAAG 451
 QY 1576 CGGCTGATTAACAGAAATTTGCTGCGCGGAGTACGCGGTGTCCTGACCTGACCCCAT 1635
 DB 452 CGGCTGATTAACAGAAATTTGCTGCGCGGAGTACGCGGTGTCCTGACCTGACCCCAT 511
 QY 1636 GCCGAACCTCAGAAAGTGAA 1653
 DB 512 GCCGAACCTCAGAAAGTGAA 529
 DB
 RESULT 12
 ABN84393
 ID ABN84393 standard; DNA; 4665 BP.
 XX
 AC ABN84393;
 XX
 DT 01-OCT-2002 (first entry)
 XX
 DE Antibiotic-independent high expression vector pHCE(III).
 XX
 KW Vector; pHCE(III); glutamate racemase; enzyme; selectable marker;
 KM gene; ds.
 XX
 OS Chimeric - Bacillus borestelensis.
 OS Chimeric - Unidentified.
 XX
 FN WO200255716-A1.
 XX
 PD 18-JUL-2002.
 XX
 PF 11-JAN-2002; 2002WO-KR00048.
 XX
 PR 12-JAN-2001; 2001KR-0001690.
 XX

PA (BIOL-) BIOLEADERS CORP.
 PA (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.
 XX
 XX Sung MH, Lee SG, Hong SP, Yoon EU, Choi YH, Poo HR;
 XX WPI; 2002-557828/59.
 DR
 XX
 XX
 PT New antibiotic-independent plasmid vector, useful for stably expressing
 PT and transcribing genes without the need for any antibiotic-resistant
 PT genes, and for producing other recombinant proteins or nucleic acids -
 XX
 XX Claim 8; Page 41-44; 45pp; English.
 PS
 CC The present sequence is that of novel antibiotic-independent high
 CC expression vector plasmid pHCE(III), deposited in Escherichia coli
 CC as KCIC 0925 BP. The vector was constructed by insertion of the
 CC Bacillus borestelensis glutamate racemase (GluA) gene (see ABN84392)
 CC into vector pHCE19T(III). The GluA gene, being essential to the
 CC microbial survival, was utilised as selectable marker. A recombinant
 CC vector can be stably transcribed and highly expressed using this
 CC vector, without the use of antibiotic resistance genes. Production
 CC of GluA from the vector is sufficient to provide D-glutamic acid
 CC for cell growth of a transformed host. The stability of the vector
 CC was confirmed using pHCE(III)-TPI, which includes a tyrosine
 CC phenol lyase gene, by transfer to E. coli WM335 (a D-glutamate
 CC auxotroph) host cells and cultivation through 9 generations. When
 CC expression vector pHCE(III)-TNA, which includes the Symbiodactylum
 CC toebii SC-1 tryptophan indole lyase (TNA) gene, was transformed
 CC into E. coli WM335 and cultivated, TNA production increased with
 CC cell growth and a high expression level was maintained throughout
 CC the growth period. The TNA content reached about 40% of total cell
 CC protein. The vector resolves problems associated with the use of
 CC expensive inducers and antibiotic-resistance genes. It can be
 CC widely applied for large-scale production of exogenous proteins,
 CC with no adverse effect on the environment.
 CC
 XX Sequence 4665 BP; 1141 A; 1193 C; 1145 G; 1186 T; 0 other;
 SQ
 Query Match 12.0%; Score 198; DB 24; Length 4665;
 Best Local Similarity 100.0%; Pred. No. 1e-40;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1456 GGAATTTCGAGCTCGGTACCCGGGGATCCTTAGAGTCGACCTGCAGCGCATGCAAGCTTGG 1515
 DB 226 GGAATTTCGAGCTCGGTACCCGGGGATCCTTAGAGTCGACCTGCAGCGCATGCAAGCTTGG 285
 QY 1516 CTGTTTGGCGGATGAGAGAAATTTTCAGCTGATACGATTAATTCGAACGCAAG 1575
 DB 286 CTGTTTGGCGGATGAGAGAAATTTTCAGCTGATACGATTAATTCGAACGCAAG 345
 QY 1576 CGGCTGATTAACAGAAATTTGCTGCGCGGAGTACGCGGTGTCCTGACCTGACCCCAT 1635
 DB 346 CGGCTGATTAACAGAAATTTGCTGCGCGGAGTACGCGGTGTCCTGACCTGACCCCAT 405
 QY 1636 GCCGAACCTCAGAAAGTGAA 1653
 DB 406 GCCGAACCTCAGAAAGTGAA 423
 DB
 RESULT 13
 ABN81604
 ID ABN81604 standard; DNA; 1506 BP.
 XX
 AC ABN81604;
 XX
 DT 29-AUG-2002 (first entry)
 XX
 DE Fungal decaprenyl diphosphate synthase encoding DNA SEQ ID NO 1.
 XX
 KW Fungi; Aspergillus; Leucosporidium; coenzyme Q10; enzyme;
 KW decaprenyl diphosphate synthase; enzyme; gene; ds.
 XX
 OS Leucosporidium scottii.
 OS

CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
 XX
 XX Sequence 1293 BP; 321 A; 375 C; 346 G; 251 T; 0 other;

Query Match 11.3%; Score 187.2; DB 23; Length 1293;
 Best Local Similarity 54.1%; Pred. No. 4e-38; Indels 21; Gaps 1;
 Matches 418; Conservative 0; Mismatches 333;

671 CAACCGCGGTGGCTGAGATCAAGAGATGATCATGACATCACTCTCCACGACGAC 730
 Db 529 CAGGACAGATGCGCTCTTTTCGAGATGAGCACTGGCCAGTGTGTCACGACGAT 588
 Qy 731 GTTATGACGCTTCCGAGACGACGAAACGACCATCCGGAACCGGATTCGGAAC 790
 Db 589 GTCATCGATCACTGCGACTTCCGACGCGGCAAGCCAGCGTGAATGCTGTGAAACAC 648
 Qy 791 AAGATGGCGATTTGGCTGCTGATTTCTTTGGGACGCGCTCTGTTCATTTGGCGAGG 850
 Db 649 AAGAAGGTCAACATGCGCTGATTAACATCTTATCGATTGCTCGATTATGATGCTCT 708
 Qy 851 TTGCGCAATCCGAGAGTATGAGCTTTTGGCTACTGTTATTCGAAACTTGGTTAGGGA 910
 Db 709 CTGCCAGCGATGATGATGACATCGTCTGATCAGATTTTACCGATTTGGTCCAAAGC 768
 Qy 911 GAGTTCATGCACTGAAAAATACTGTTGATGATGCGATTAGGCTACGCGGACGAGGAA 970
 Db 769 GAGTTCATGCACTGAGGCTCAAGGAAACGGA-----GAAACGAG 807
 Qy 971 ACGTTCGATTAATTTTGGAGAAAGACTTCTGAAGACTGCGCTCTTGAATTCGCAAGTCG 1030
 Db 808 CGCTTCGCCCATTAATCTGACCAAGACATACAGAAAGCCGATCGTGTATCCCAATGCA 867
 Qy 1031 TGCAAGCAAGTGGCTTCTGAGTGTGTAGCCCTGAGAGTGTGATCTCTTATGCT 1090
 Db 868 CTGAAGCGGACCGCGCTGATTTCCAGCGACGACACAGTGGCCGAGTGGCTTCCAG 927
 Qy 1091 TACGGAAGAACTTGTGTTGTCATTCAGATCGTGCAGCAGATGCTGCATACACCGTC 1150
 Db 928 TACGGAAGCAATCGGCTTGGCTTTGAGTGTGTCAGATGATGCTGGAATTTGCTCC 987
 Qy 1151 TCCGTTACCGACTCGTGAAGCCCGCGGTGCAAGCTTCAAGTGGTTCGCCACCGG 1210
 Db 988 TCCACCGAGCAAGATGGGCAAGCCGAGCGGCGGATTTGAAGCTGGTCTGCCACCGCT 1047
 Qy 1211 CCGGCGCTCTTGGCATGGAAGACCAACCGCGAGCTCGGCTCCCATGATCAAGGCAAGTTC 1270
 Db 1048 CCGGCTCTCTTTGCAATGCGAAGATACCCCGAGCTGAATCCATGTGATGCGCGCTTC 1107
 Qy 1271 TCTGACCCGAGAGAGCTGAGCGTGCACGCGATTGTCGAAAAAGTATGATTTGAG 1330
 Db 1108 AGCGAGCCCGGAGAGCTGAGCGCGCTTCGAGCTGTGTCACAAAGCGACGCTGTGAG 1167
 Qy 1331 AAGACGAGAGCTTGGCGGAGAGATGCGCAAGAGGCTTGGATGCAATTTGGACGTTTC 1390
 Db 1168 CAGACCGCGTTTCTGCGCCAAAGACACTGCAACGAGGCGATTCGCGTGGCCACGAGAGCTC 1227
 Qy 1391 CCGGAGATCCGCGACGGAAGGCTTGGAGCATTTGAGCGGACGGAAGGTTGTA 1442
 Db 1228 ACGAGTGCCTTACCAAGAGGCTTCAGGTGTGCGGACTTGTATCA 1279

RESULT 15
 AAC61900
 ID AAC61900 standard; DNA; 966 BP.
 XX
 AC AAC61900;

XX 06-MAR-2001 (first entry)
 DT DNA encoding straight-chain class prenyltransferase designated ATP8.
 XX
 XX
 DE
 XX Prenyltransferase; ATP1; ATP2; ATP3; ATP4; ATP5; ATP6; ATP7;
 KW ATP8; ATP9; ATP10; ATP11; ATP12; tocopherol; homogentisic acid;
 KW phytylpyrophosphate; 2-methyl-6phytylbenzoquinol; antioxidant;
 KW nutritional supplement; ss.
 XX
 OS Arabidopsis sp.
 XX
 XX
 XX Key Location/Qualifiers
 FT 1..966
 FT CDS /*tag= a
 FT /product= "prenyltransferase ATP8"

MO200063391-A2.

26-OCT-2000.

14-APR-2000; 2000MO-US10368.

15-APR-1999; 99US-0129999.
 30-JUL-1999; 99US-0146461.

(CALJ) CALGENE LLC.

Savidge B, Lassner MW, Weiss JD, Post-Beltemmiller D;

WPI; 2000-647519/62.

P-PSDB; AAB19412.

An isolated nucleic acid sequence encoding prenyltransferase used to transform plant cells to increase the production of tocopherols -

Example 1; Page 82; 114pp; English.

The present sequence encodes a prenyltransferase. The specification describes prenyltransferases designated ATP1, ATP2, ATP3, ATP4, ATP5, ATP6, ATP7, ATP8, ATP9, ATP10, ATP11, and ATP12. The biosynthesis of alpha-tocopherol in higher plants involves the condensation of homogentisic acid and phytylpyrophosphate to form 2-methyl-6phytylbenzoquinol, which can form various tocopherols. The prenyltransferase polynucleotides are useful in transforming host cells to alter the expression of prenyltransferase in these cells. The transformed cells are used in the production of tocopherols which are of use in the pharmaceutical industry as antioxidants and also in the food industry as nutritional supplements.

Sequence 966 BP; 278 A; 196 C; 240 G; 252 T; 0 other;

Query Match 10.9%; Score 180.8; DB 21; Length 966;
 Best Local Similarity 54.1%; Pred. No. 1.ee-36;
 Matches 405; Conservative 0; Mismatches 322; Indels 21; Gaps 1;

650 ACGTGAATATCTCGCTCGCAACGCGGCTGCTGATCAGATCAGAGATTCATGCA 709
 Db 163 ACATCAGATTACCGGTAAAGCAAGCGGTATGCTGAATCACTGAATATATACGTC 222
 Qy 710 GCATCACTCTCCACGACGACGATATGACGCTTCCGAGACCAAGCAAGCAATCC 769
 Db 223 GCAAGTCTACTGACAGATGATGCTTGGATATGCCATACAAAGCGGCTGTGTTCC 282
 Qy 770 GGAACACGAGATTCGAAAGAAAGATGCGATTTGGCTGTGATTTCTGTTGGAGCG 829
 Db 283 TTAATGTTGTAATGAGTAAACAGATGCGGTATTAAGAGAGACTTCTCTCCCGG 342
 Qy GCGTCTGTTGATTTGGAGAGGTTGCGCAATCCGAGGATTTGACTTTGGTACTGTT 889
 Db 343 GCTTGTGGGCTCTCGCTCTTTAAAGAACACAGAGTTGACATTAATCTGCAACTGCT 402
 Qy 890 ATTGAAACTTGTGTTAGGAGGAGAGTTCATGCAATTGAAAAAACTGTTGATGATTCGATT 949

Db 403 GTAGACATCTTTGTTACCGGTGAACCATGGAATACTAGT----- 444
QY 950 GAGGCTACGGCGACGCGAGAAACGTTTCGATTACTATTTCGAGAAGACTTACTTTGAAGACT 1009
Db 445 ---TCAACCGAGCAGCGGTATAGTATGACTACTACATGCAGAAGACATATTATAAGACA 501
QY 1010 GCGTCTTGTATTGCCAAGTCGTGCAGAGCAAGTCGCTTCTGGGTGGTGTACGCTGAG 1069
Db 502 GCATCGCTAATCTTAACAGCTGCAAGCTGTTGCCGTTCTCACTGGACAACACAGCAGAA 561
QY 1070 GTTGCTGATCGTCTATGCTTACGGAAGGAACCTTGGTTTGGCATTTCAGATCGTCGAC 1129
Db 562 GTTGCGGTGTAGCTTTTGAGTATGGAGGAATCTGGGTTTAGCATTCCAATTAATAGAC 621
QY 1130 GACATGCTGACTACACCGTCTCCGTACCGACTCGGTAAGCCCGCGGTGCAGACCTC 1189
Db 622 GACATTCTTGATTTCACGGGCACATCTGCCTCTCTCGGAAGGGATCGTTGTTCAGATATT 681
QY 1190 CAGCTCGGTCTGCCACCGCGCGCCCTCTTCGCATGGGAAGCACACCGCGAGCTCGGT 1249
Db 682 CGCCATGGAGTCAATAACAGCCCCAATCCTCTTTGCCATGGAGAGTTTCTCAACTACGC 741
QY 1250 CCCATGATCAAGCGCAAGTTCTCTGACCCAGAGAGCGTCGAGCGTGCACGCGAGTTGGTC 1309
Db 742 GAAGTTGTTGATCAAGTTGAAAGATCCTAGGAATGTTGACATTGCTTTAGAGTATCTT 801
QY 1310 GAGAAAAGTGATGGATTGGAGAAAGACGAGAGCCCTTGGCGGAGGAGTATGCCCAAGAGCGG 1369
Db 802 GGAAGAGCAAGGGAATACAGAGGCAAGAGAATTAGCCATGGAACATGGGAATCTAGCA 861
QY 1370 TTGGATGCAATTCGGAGCTTCCCGGAGA 1397
Db 862 GCAGCTGCAATCGGCTCTCTACCTGAAA 889

Search completed: January 16, 2003, 08:23:14
Job time : 277 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 07:27:51 ; Search time 58 Seconds

(without alignments)
8740.288 Million cell updates/sec

Title: US-09-830-111A-1
Perfect score: 1653
Sequence: 1 ttttggggtcgaaagtc.....atgcggaactcagaagtga 1653

Scoring table: IDENTITY NUC

Gapop 10.0 , Capext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	179.2	10.8	5926	4	US-09-027-169-3
C 2	179.2	10.8	5926	4	US-09-027-169-4
C 3	160.4	9.7	3301	2	US-08-447-430A-42
4	155.2	9.4	4009	2	US-08-500-860A-2
5	152.4	9.2	3423	2	US-08-447-430A-40
6	152.4	9.2	3423	4	US-09-485-737B-86
7	152	9.2	1097	1	US-08-761-258-7
8	152	9.2	1097	2	US-08-977-306-7
9	148.8	9.0	6312	1	US-08-531-601-3
10	148.8	9.0	6312	2	US-08-859-032-3
11	147.2	8.9	518	4	US-09-242-901-3
12	147	8.9	3474	2	US-08-447-430A-41
13	147	8.9	3474	2	US-08-318-837-10
14	147	8.9	3474	4	US-08-122-458B-8
C 15	147	8.9	8430	4	US-09-131-028A-6
16	147	8.9	8430	4	US-09-131-028A-10
17	146.2	8.8	628	4	US-09-242-901-1
18	146	8.8	4476	3	US-08-801-344-2
19	146	8.8	4476	4	US-09-498-599-2
20	145.6	8.8	7475	2	US-08-971-036-1
21	145.6	8.8	7475	3	US-09-096-570-1
22	145.6	8.8	7475	4	US-09-265-617B-1
C 23	145.4	8.8	1200	4	US-09-204-117B-2
C 24	145.4	8.8	2650	4	US-09-204-117B-3
C 25	145.4	8.8	4199	4	US-09-204-117B-1
C 26	144	8.7	5096	1	US-08-106-433A-3
27	144	8.7	5110	1	US-08-106-433A-4

28	144	8.7	5110	1	US-08-106-433A-5	Sequence 5, Appli
29	144	8.7	5110	1	US-08-106-433A-6	Sequence 6, Appli
30	144	8.7	5110	1	US-08-106-433A-7	Sequence 7, Appli
31	144	8.7	5110	1	US-08-106-433A-8	Sequence 8, Appli
32	144	8.7	5110	1	US-08-106-433A-9	Sequence 9, Appli
33	144	8.7	5110	1	US-08-106-433A-10	Sequence 10, Appli
34	144	8.7	5110	1	US-08-106-433A-11	Sequence 11, Appli
35	144	8.7	5110	1	US-08-106-433A-12	Sequence 12, Appli
36	144	8.7	5110	1	US-08-106-433A-13	Sequence 13, Appli
37	144	8.7	5110	1	US-08-106-433A-14	Sequence 14, Appli
38	144	8.7	5110	1	US-08-106-433A-15	Sequence 15, Appli
39	138	8.3	5238	6	5453363-1	Patent No. 5453363
C 40	135.4	8.2	4593	3	US-08-801-344-1	Sequence 1, Appli
C 41	135.4	8.2	4593	4	US-09-498-599-1	Sequence 1, Appli
42	129.8	7.9	469	2	US-08-647-368A-1	Sequence 18, Appli
43	120.4	7.3	1464	4	US-09-276-531-18	Sequence 1, Appli
44	92	5.6	2263	2	US-08-687-865A-1	Sequence 1, Appli
45	92	5.6	2263	4	US-09-043-711-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-027-169-3/c
; Sequence 3, Application US/09027169
; Patent No. 6420524
; GENERAL INFORMATION:
; APPLICANT: CRAIG, NANCY L
; TITLE OF INVENTION: GAIN OF FUNCTION MUTATIONS IN
; NUMBER OF INVENTION: ATP-DEPENDENT TRANSPOSITION PROTEINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Anne Brown (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave.
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; ZIP: 27608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/027,169
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Anne
; REGISTRATION NUMBER: 36,463
; REFERENCE/DOCKET NUMBER: 5789-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2205
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5926 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "PEM delta R.adj to 1"
US-09-027-169-3

Query Match 10.8%; Score 179.2; DB 4; Length 5926;
Best Local Similarity 91.3%; Pred. No. 1.2e-37;
Matches 190; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1446 GGTCAAGATAGGAATTCGAGTCGGTACCGGGGATCCTCTAGAGTCGACCTCGAGGCAT 1505
Db 4282 GTTATTAAAGCTGTCGACGGACTTCTTTGGGGATCCTCTAGAGTCGACCTCGAGGCAT 4223

QY 1506 GCAAGCTTGCTGTTTGGCGGATGAGAGATTTTACCTGATACAGATTAAATCAG 1565
DB 4222 GCAAGCTTGCTGTTTGGCGGATGAGAGATTTTACCTGATACAGATTAAATCAG 4163
QY 1566 AACGCAAGACGGCTGTGATTAACAGAAATTTGCTGGCGGACAGTACGGCGGTGCTCCAC 1625
DB 4162 AACGCAAGACGGCTGTGATTAACAGAAATTTGCTGGCGGACAGTACGGCGGTGCTCCAC 4103
QY 1626 CTGACCCCATGCCGAAGTCAAGAGTGA 1653
DB 4102 CTGACCCCATGCCGAAGTCAAGAGTGA 4075

RESULT 2
US-09-027-169-4/c
Sequence 4, Application US/09027169
Patent No. 6420524
GENERAL INFORMATION:
APPLICANT: CRAIG, NANCY L
TITLE OF INVENTION: GAIN OF FUNCTION MUTATIONS IN
TITLE OF INVENTION: ATP-DEPENDENT TRANSDUCTION PROTEINS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Anne Brown (Alston & Bird, LLP)
STREET: 3605 Glenwood Ave.
CITY: Raleigh
STATE: NC
COUNTRY: USA
ZIP: 27608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/027,169
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Anne
REGISTRATION NUMBER: 36,463
REFERENCE/DOCKET NUMBER: 5789-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919 420 2205
TELEFAX: 919 881 3175
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 5926 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: circular
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "pBW-delta"
US-09-027-169-4

Query Match 10.8%; Score 179.2; DB 4; Length 5926;
Best Local Similarity 91.3%; Pred. No. 1.2e-37;
Matches 190; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1446 GGTCAAGATGGAATTTGAGCTGCGTACCCGGGATCTCTAGAGTGCAGCTGAGGCAT 1505
DB 3918 GTTATATTAAGCTGTCCGACGAGCTTCTTGGGGATCTCTAGAGTGCAGCTGAGGCAT 3859
QY 1506 GCAAGCTTGCTGTTTGGCGGATGAGAGATTTTACCTGATACAGATTAAATCAG 1565
DB 3858 GCAAGCTTGCTGTTTGGCGGATGAGAGATTTTACCTGATACAGATTAAATCAG 3799
QY 1566 AACGCAAGACGGCTGTGATTAACAGAAATTTGCTGGCGGACAGTACGGCGGTGCTCCAC 1625
DB 3798 AACGCAAGACGGCTGTGATTAACAGAAATTTGCTGGCGGACAGTACGGCGGTGCTCCAC 3739
QY 1626 CTGACCCCATGCCGAAGTCAAGAGTGA 1653

DB 3738 CTGACCCCATGCCGAAGTCAAGAGTGA 3711

RESULT 3
US-08-447-430A-42
Sequence 42, Application US/08447430A
Patent No. 5916558
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Recombinant polypeptides and peptides,
TITLE OF INVENTION: nucleic acids coding for the same and use of these
TITLE OF INVENTION: polypeptides and peptides in the diagnostic of
TITLE OF INVENTION: tuberculosis.
NUMBER OF SEQUENCES: 43
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,430A
FILING DATE:
CLASSIFICATION: 424
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 3301 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: circular
MOLECULE TYPE: plasmid vector
HYPOTHETICAL: NO
US-08-447-430A-42

Query Match 9.7%; Score 160.4; DB 2; Length 3301;
Best Local Similarity 97.8%; Pred. No. 9e-33;
Matches 174; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1476 GGGAGCTCTGAGCTGCACTGCGAGCATGCAAGCTTGGCTGTTTGGCGGATGAGAGA 1535
DB 237 GGGAGCTCTGAGCTGCACTGCGAGCATGCAAGCTTGGCTGTTTGGCGGATGAGAGA 293
QY 1536 AGATTTTGCCTGATGATGATTAATCAGAACGAGAGCGGTGTGATTAACAGATT 1595
DB 294 AGATTTTGCCTGATGATGATTAATCAGAACGAGAGCGGTGTGATTAACAGATT 353
QY 1596 TGCCTGCGGCGAGTACGGCGGTGTCCACAGTACCCCATGCCGAAGTGA 1653
DB 354 TGCCTGCGGCGAGTACGGCGGTGTCCACAGTACCCCATGCCGAAGTGA 411

RESULT 4
US-08-500-860A-2
Sequence 2, Application US/08500860A
Patent No. 5891679
GENERAL INFORMATION:
APPLICANT: LUCAS, RUDOLPH
APPLICANT: DE BAETSELIER, PATRICK
APPLICANT: FRANSSEN, LUCIE
APPLICANT: SABON, ERWIN
TITLE OF INVENTION: TNF-ALPHA, A PROCESS FOR PREPARING THEM AND
TITLE OF INVENTION: THEIR USE AS ACTIVE SUBSTANCES IN PHARMACEUTICAL COMPOSITIONS
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/500,860A

FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.

REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-8

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4000
TELEFAX: (703)816-4100

TELEX: 200797 NIXN UR
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 4009 base pairs

TYPE: nucleic acid
STRAINEDNESS: single

TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO
ANTI-SENSE: NO

IMMEDIATE SOURCE:
CLONE: pIG2

US-08-500-860A-2

Query Match 9.4%; Score 155.2; DB 2; Length 4009;
Best Local Similarity 96.6%; Pred. No. 2.3e-31;
Matches 170; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 1478 GGATCCTAGAGTCGACCTGCGGATGCAAGCTTGGCTGTTTGGCGGATGAGAG 1537

Db 947 GAATTCCTAGAGTCGACCTGCGGATGCAAGCTTGGCTGTTTGGCGGATGAGAG 1003

QY 1538 ATTTTCAGCTGTATACAGATTAATCAGAACGCGAGCGGCTCTGATAAAACAGAAATTG 1597

Db 1004 ATTTTCAGCTGTATACAGATTAATCAGAACGCGAGCGGCTCTGATAAAACAGAAATTG 1063

QY 1598 CTTGGCGGAGTAGGCGGTGGTCCCACTGACCCCATGCCAACTCAGAAAGTGA 1653

Db 1064 CTTGGCGGAGTAGGCGGTGGTCCCACTGACCCCATGCCAACTCAGAAAGTGA 1119

RESULT 5
US-08-447-430A-40

Sequence 40, Application US/08447430A
Patent No. 5916558

GENERAL INFORMATION:
APPLICANT:

TITLE OF INVENTION: Recombinant polypeptides and peptides,
nucleic acids coding for the same and use of these

TITLE OF INVENTION: polypeptides and peptides in the diagnostic of
tuberculosis.

NUMBER OF SEQUENCES: 43
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,430A

FILING DATE:
CLASSIFICATION: 424

INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:

LENGTH: 3423 base pairs
TYPE: nucleic acid

STRAINEDNESS: single
TOPOLOGY: circular

MOLECULE TYPE: plasmid vector
HYPOTHETICAL: NO

US-08-447-430A-40

Query Match 9.2%; Score 152.4; DB 2; Length 3423;
Best Local Similarity 97.6%; Pred. No. 1.2e-30;
Matches 166; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1484 TCTAGAGTCGACCTGCGGATGCAAGCTTGGCTGTTTGGCGGATGAGAGATTTC 1543

Db 363 TCTAGAGTCGACCTGCGGATGCAAGCTTGGCTGTTTGGCGGATGAGAGATTTC 419

QY 1544 AGCTGTATACAGATTAATCAGAACGCGGCTCTGATAAAACAGAAATTTCCTGCG 1603

Db 420 AGCTGTATACAGATTAATCAGAACGCGGCTCTGATAAAACAGAAATTTCCTGCG 479

QY 1604 GGCAGTAGCGCGGTGGTCCCACTGACCCCATGCCAACTCAGAAAGTGA 1653

Db 480 GGCAGTAGCGCGGTGGTCCCACTGACCCCATGCCAACTCAGAAAGTGA 529

RESULT 6
US-09-485-737B-86

Sequence 86, Application US/09485737B
Patent No. 6350860

GENERAL INFORMATION:
APPLICANT: Buyse, Marie-Ange

APPLICANT: Sablon, Erwin
TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC SHOCK,

CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS
TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS

FILE REFERENCE: INNS:015
CURRENT APPLICATION NUMBER: US/09/485,737B

PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: PCT/EP 98/05165

PRIOR FILING DATE: 1998-08-14
PRIOR APPLICATION NUMBER: EPO 98870139.7

PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: EPO 97870122.5

PRIOR FILING DATE: 1997-08-18
NUMBER OF SEQ ID NOS: 104

SOFTWARE: PatentIn version 3.0
SEQ ID NO 86

LENGTH: 3423
TYPE: DNA

ORGANISM: UNKNOWN
FEATURE:

OTHER INFORMATION: GENOMIC
US-09-485-737B-86

Query Match 9.2%; Score 152.4; DB 4; Length 3423;
Best Local Similarity 97.6%; Pred. No. 1.2e-30;
Matches 166; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1484 TCTAGAGTCGACCTGCGGATGCAAGCTTGGCTGTTTGGCGGATGAGAGATTTC 1543

Db 363 TCTAGAGTCGACCTGCGGATGCAAGCTTGGCTGTTTGGCGGATGAGAGATTTC 419

QY 1544 AGCTGTATACAGATTAATCAGAACGCGGCTCTGATAAAACAGAAATTTCCTGCG 1603

Db 420 AGCTGTATACAGATTAATCAGAACGCGGCTCTGATAAAACAGAAATTTCCTGCG 479

QY 1604 GGCAGTAGCGCGGTGGTCCCACTGACCCCATGCCAACTCAGAAAGTGA 1653

Db 480 GGCAGTAGCGCGGTGGTCCCACTGACCCCATGCCAACTCAGAAAGTGA 529

RESULT 7
US-08-761-258-7

Sequence 7, Application US/08761258
Patent No. 5756087

GENERAL INFORMATION:
APPLICANT: Ligon, James M.

APPLICANT: Hill, Dwight S.
APPLICANT: Lam, Stephen T.

APPLICANT: Gaffney, Thomas D.
APPLICANT: Torkewitz, Nancy

TITLE OF INVENTION: Genetically Modified Pseudomonas Strains

TITLE OF INVENTION: With Enhanced Biocontrol Activity
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ciba-Geigy Corporation
 STREET: 520 White Plains Road, P.O. Box 2005
 CITY: Tarrytown
 STATE: NY
 COUNTRY: USA
 ZIP: 10591
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/761,258
 FILING DATE:
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Weigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (919) 541-8587
 TELEFAX: (919) 541-8689
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1097 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 IMMEDIATE SOURCE:
 CLONE: pK223-3
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 1
 OTHER INFORMATION: /note= "BssHII site"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 7
 OTHER INFORMATION: /note= "BglII site"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 265
 OTHER INFORMATION: /note= "EcoRI site"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 270
 OTHER INFORMATION: /note= "SmaI site"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 275
 OTHER INFORMATION: /note= "BamHI site"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 281
 OTHER INFORMATION: /note= "SalI site"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 287
 OTHER INFORMATION: /note= "PstI site"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 293
 OTHER INFORMATION: /note= "XbaI site"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 299
 OTHER INFORMATION: /note= "XhoI site"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 305
 OTHER INFORMATION: /note= "KpnI site"

FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 311
 OTHER INFORMATION: /note= "No. 57560871 site"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 319
 OTHER INFORMATION: /note= "HindIII site"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 1086
 OTHER INFORMATION: /note= "BglI site"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 1092
 OTHER INFORMATION: /note= "BssHII site"
 FEATURE:
 NAME/KEY: promoter
 LOCATION: 185..264
 OTHER INFORMATION: /standard_name= "lac"
 FEATURE:
 NAME/KEY: terminator
 LOCATION: 327..752
 OTHER INFORMATION: /standard_name= "trnB"
 US-08-761-258-7
 Query Match 9.2%; Score 152; DB 1; Length 1097;
 Best Local Similarity 77.1%; Pred. No. 1e-30;
 Matches 185; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 QY 1414 TTGTGACAGTGGACGACAAAGCTTTGACTAGTCAGTCAAGATGAGAAATTCGGCTGAC 1473
 DB 225 TGTGAAATTGTGAGGATTAACAATTTCCACAGGAAACAGAAATTCCTGGGATCCGTCG 284
 QY 1474 CCGGGATCCTCTAGAGTGCACCTGCAGCATGCAAGCTTGCTTTTGGCGATGAGA 1533
 DB 285 ACCTGACGTCTGACTCGAGGTACCGGCGCGAAGCTTGGCTTTTGGCGATGAGA 344
 QY 1534 GAAGATTTTCAGCTGATACAGATTAAATCAAGAAAGCAAGAAAGCGTCTGATTAACAGAA 1593
 DB 345 GAAGATTTTCAGCTGATACAGATTAAATCAAGAAAGCAAGAAAGCGTCTGATTAACAGAA 404
 QY 1594 TTGGCTGCGCGACATGACGGGGTGTCCACCTGACCCCAATGCCGAATCAGAAAGTAA 1653
 DB 405 TTGGCTGCGCGACATGACGGGGTGTCCACCTGACCCCAATGCCGAATCAGAAAGTAA 464
 RESULT 8
 US-08-977-306-7
 Sequence 7, Application US/08977306
 Patent No. 5955348
 GENERAL INFORMATION:
 APPLICANT: Ligon, James M.
 APPLICANT: Hill, Dwight S.
 APPLICANT: Gaffney, Thomas D.
 APPLICANT: Torkewitz, Nancy
 APPLICANT: Stafford, Jill M.
 TITLE OF INVENTION: Genetically Modified Pseudomonas Strains
 TITLE OF INVENTION: With Enhanced Biocontrol Activity
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 5955348artis Corporation
 STREET: 3054 Cornwallis Road
 CITY: Research Triangle Park
 STATE: NC
 COUNTRY: USA
 ZIP: 27709
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

```
;; APPLICATION NUMBER: US/08/977,306
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meigs, J. Timothy
;; REGISTRATION NUMBER: 38,241
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (919) 541-8587
;; TELEFAX: (919) 541-8689
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1097 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; IMMEDIATE SOURCE:
;; CLONE: pKK223-3
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 1
;; OTHER INFORMATION: /note= "BssHII site"
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 7
;; OTHER INFORMATION: /note= "BglII site"
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 265
;; OTHER INFORMATION: /note= "EcoRI site"
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 270
;; OTHER INFORMATION: /note= "SmaI site"
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 275
;; OTHER INFORMATION: /note= "BamHI site"
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 281
;; OTHER INFORMATION: /note= "SalI site"
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 287
;; OTHER INFORMATION: /note= "PstI site"
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 293
;; OTHER INFORMATION: /note= "XbaI site"
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 299
;; OTHER INFORMATION: /note= "XhoI site"
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 305
;; OTHER INFORMATION: /note= "KpnI site"
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 311
;; OTHER INFORMATION: /note= "No. 5955348I site"
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 319
;; OTHER INFORMATION: /note= "HindIII site"
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 1086
;; OTHER INFORMATION: /note= "BglI site"
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 1092
```

```
;; OTHER INFORMATION: /note= "BssHII site"
;; FEATURE:
;; NAME/KEY: promoter
;; LOCATION: 185..264
;; OTHER INFORMATION: /standard_name= "tac"
;; FEATURE:
;; NAME/KEY: terminator
;; LOCATION: 327..752
;; OTHER INFORMATION: /standard_name= "rrnB"
US-08-977-306-7

Query Match          9.2%; Score 152; DB 2; Length 1097;
Best Local Similarity 77.1%; Pred No. 1e-30;
Matches 185; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1414 TTGTGGAGCAGTTGACGCAAGGTGTTGACTAGTCAAGATAGGAATTCGAGTCGGTAC 1473
DB 225 TGTGGAATTGTGAGCGGATAACAATTTACACAGGAACAGAAATCCCGGGGATCCGTCG 284
QY 1474 CCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCTGTTTGGCGGATGAGA 1533
DB 285 ACCTGCAGTCTAGACTCGAGGTACCGCGCGCAAGCTTGGCTGTTTGGCGGATGAGA 344
QY 1534 GAAGATTTTCAGCCTCATACAGATTAAATCAGAACGCAAGCGGTCTGTATAAACAGAA 1593
DB 345 GAAGATTTTCAGCCTCATACAGATTAAATCAGAACGCAAGCGGTCTGTATAAACAGAA 404
QY 1594 TTTCCTGCGCGCAGTAGCGCGGTGTCACCTGACCCCATGCCGAACCTCAGAAGTGAA 1653
DB 405 TTTCCTGCGCGCAGTAGCGCGGTGTCACCTGACCCCATGCCGAACCTCAGAAGTGAA 464

RESULT 9
US-08-531-601-3
; Sequence 3, Application US/08531601
; Patent No. 5686684
; GENERAL INFORMATION:
; APPLICANT: YOSHIGI, NAOHIRO
; APPLICANT: MAEBA, HIDEO
; APPLICANT: OKADA, YUKIO
; TITLE OF INVENTION: RECOMBINANT BETA-AMYLASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,601
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP HEI 6-233086
; FILING DATE: 28-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 08/531,601
; REFERENCE/DOCKET NUMBER: 2589-031-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6312 base pairs
; TYPE: nucleic acid
```

STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-531-601-3

Query Match 9.0%; Score 148.8; DB 1; Length 6312;
Best Local Similarity 98.7%; Pred. No. 1.3e-29;
Matches 150; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1502 GCATGCAAGCTTGCTGTTGGCGGATGAGAGAAGATTTTCAGCTGATACAGATTAA 1561
DB 1747 GCAGCAAGCTTGCTGTTGGCGGATGAGAGAAGATTTTCAGCTGATACAGATTAA 1806
QY 1562 TCAGAACCGAAGAGCGGTCTGTATAAACAGATTTCCTGGCGGAGTACCGCGTGTGTC 1621
DB 1807 TCAGAACCGAAGAGCGGTCTGTATAAACAGATTTCCTGGCGGAGTACCGCGTGTGTC 1866
QY 1622 CCACCTGACCCCATGCCGAACTCAGAAAGTGAA 1653
DB 1867 CCACCTGACCCCATGCCGAACTCAGAAAGTGAA 1898

RESULT 10
US-08-859-032-3
Sequence 3, Application US/08859032

Patent No. 5863784

GENERAL INFORMATION:

APPLICANT: YOSHIGI, NAOKIHO

APPLICANT: MAEDA, HIDEO

APPLICANT: OKADA, YUKIO

TITLE OF INVENTION: RECOMBINANT BETA-AMYLASE

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT

STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/859,032

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/531,601

FILING DATE: 21-SEP-1995

APPLICATION NUMBER: JP HEI 6-233086

FILING DATE: 28-SEP-1994

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F

REGISTRATION NUMBER: 08/531,601

REFERENCE/DOCKET NUMBER: 2589-031-0

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 6312 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

US-08-859-032-3

Query Match 9.0%; Score 148.8; DB 2; Length 6312;

Best Local Similarity 98.7%; Pred. No. 1.3e-29;

Matches 150; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1502 GCATGCAAGCTTGCTGTTGGCGGATGAGAGAAGATTTTCAGCTGATACAGATTAA 1561
DB 1747 GCAGCAAGCTTGCTGTTGGCGGATGAGAGAAGATTTTCAGCTGATACAGATTAA 1806
QY 1562 TCAGAACCGAAGAGCGGTCTGTATAAACAGATTTCCTGGCGGAGTACCGCGTGTGTC 1621
DB 1807 TCAGAACCGAAGAGCGGTCTGTATAAACAGATTTCCTGGCGGAGTACCGCGTGTGTC 1866
QY 1622 CCACCTGACCCCATGCCGAACTCAGAAAGTGAA 1653
DB 1867 CCACCTGACCCCATGCCGAACTCAGAAAGTGAA 1898

RESULT 11
US-09-242-901-3
Sequence 3, Application US/09242901

Patent No. 6322783

GENERAL INFORMATION:

APPLICANT: TAKAHASHI, SEISHI

TITLE OF INVENTION: NOVEL BACTERIOPHAGES, METHOD FOR SCREENING SAME AND

TITLE OF INVENTION: BACTERICIDAL COMPOSITIONS USING SAME, AND DETECTION

TITLE OF INVENTION: KITS USING SAME

FILE REFERENCE: 3000-0001

CURRENT APPLICATION NUMBER: US/09/242,901

CURRENT FILING DATE: 1999-02-26

EARLIER APPLICATION NUMBER: PCT/JP97/02957

EARLIER FILING DATE: 1997-08-26

EARLIER APPLICATION NUMBER: 8/261132

EARLIER FILING DATE: 1996-08-26

EARLIER APPLICATION NUMBER: 9/135716

EARLIER FILING DATE: 1997-04-19

NUMBER OF SEQ ID NOS: 24

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 3

LENGTH: 518

TYPE: DNA

ORGANISM: Bacteriophage

FEATURE:

OTHER INFORMATION: "n" at various positions throughout the sequence may be

OTHER INFORMATION: a, t, g, c other or unknown

US-09-242-901-3

Query Match 8.9%; Score 147.2; DB 4; Length 518;
Best Local Similarity 98.0%; Pred. No. 1.4e-29;
Matches 149; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1502 GCATGCAAGCTTGCTGTTGGCGGATGAGAGAAGATTTTCAGCTGATACAGATTAA 1561
DB 189 GAAAGAAAGCTTGCTGTTGGCGGATGAGAGAAGATTTTCAGCTGATACAGATTAA 248
QY 1562 TCAGAACCGAAGAGCGGTCTGTATAAACAGATTTCCTGGCGGAGTACCGCGTGTGTC 1621
DB 249 TCAGAACCGAAGAGCGGTCTGTATAAACAGATTTCCTGGCGGAGTACCGCGTGTGTC 308
QY 1622 CCACCTGACCCCATGCCGAACTCAGAAAGTGAA 1653
DB 309 CCACCTGACCCCATGCCGAACTCAGAAAGTGAA 340

RESULT 12
US-08-447-430A-41
Sequence 41, Application US/08447430A

Patent No. 5916558

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Recombinant polypeptides and peptides,

TITLE OF INVENTION: nucleic acids coding for the same and use of these

TITLE OF INVENTION: polypeptides and peptides in the diagnosis of

TITLE OF INVENTION: tuberculosis.

NUMBER OF SEQUENCES: 43

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,430A
; FILING DATE:
; CLASSIFICATION: 424
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: plasmid vector
; HYPOTHETICAL: NO
; US-08-447-430A-41

Query Match      8.9%; Score 147; DB 2; Length 3474;
Best Local Similarity 100.0%; Pred. No. 3.1e-29;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1507 CAAGCTTGGCTGTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGA 1566
Db 436 CAAGCTTGGCTGTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGA 495
QY 1567 ACGCAGAAGCGGTCTGATAAAACAGAAATTTGCCTGGCGGAGTAGCGCGTGTCCACC 1626
Db 496 ACGCAGAAGCGGTCTGATAAAACAGAAATTTGCCTGGCGGAGTAGCGCGTGTCCACC 555
QY 1627 TGACCCCATGCCGAATCTCAGAAGTGAA 1653
Db 556 TGACCCCATGCCGAATCTCAGAAGTGAA 582

RESULT 13
US-08-318-837-10
; Sequence 10, Application US/08318837
; Patent No. 5981277
; GENERAL INFORMATION:
; APPLICANT: FRANSEN, LUCIA; DEVOS, KATHLEEN; VAN DE VOORDE,
; APPLICANT: ANDRE; VAN HEUVERSWYN, HUGO
; TITLE OF INVENTION: NEW POLYPEPTIDES AND PEPTIDES, NUCLEIC ACID
; TITLE OF INVENTION: CODING FOR THEM, AND THEIR USE IN THE FIELD OF TUMOR THERAPY
; TITLE OF INVENTION: IMMUNOLOGY
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN AND MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,837
; FILING DATE: 13-OCT-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP 93/01022
; FILING DATE: 28-APR-1993
; PRIOR APPLICATION NUMBER: 92.401.231.3
; FILING DATE: 30-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 410.007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000

```

```

; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: pMTNF-MPH
; US-08-318-837-10

Query Match      8.9%; Score 147; DB 2; Length 3474;
Best Local Similarity 100.0%; Pred. No. 3.1e-29;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1507 CAAGCTTGGCTGTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGA 1566
Db 436 CAAGCTTGGCTGTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGA 495
QY 1567 ACGCAGAAGCGGTCTGATAAAACAGAAATTTGCCTGGCGGAGTAGCGCGTGTCCACC 1626
Db 496 ACGCAGAAGCGGTCTGATAAAACAGAAATTTGCCTGGCGGAGTAGCGCGTGTCCACC 555
QY 1627 TGACCCCATGCCGAATCTCAGAAGTGAA 1653
Db 556 TGACCCCATGCCGAATCTCAGAAGTGAA 582

RESULT 14
US-08-122-458D-8
; Sequence 8, Application US/08122458D
; Patent No. 6387372
; GENERAL INFORMATION:
; APPLICANT: Cocito, Carlo
; APPLICANT: Coene, Marc
; APPLICANT: DeKessel, Myrian
; APPLICANT: Gilot, Philippe
; TITLE OF INVENTION: Polypeptides from Mycobacterium
; TITLE OF INVENTION: Paratuberculosis
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 3100 No. 6387372west Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/122,458D
; FILING DATE: 24-SEP-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP92/00661
; FILING DATE: 24-MAR-1992
; PRIOR APPLICATION NUMBER: EP 91400798.4
; FILING DATE: 25-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kowalchuk, Katherine M.
; REGISTRATION NUMBER: 36,848
; REFERENCE/DOCKET NUMBER: 8076.111USWO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; INFORMATION FOR SEQ ID NO: 8:

```

```

; SEQUENCE CHARACTERISTICS:
;   LENGTH: 3474 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-08-122-458D-8

```

```

Query Match      8.9%; Score 147; DB 4; Length 3474;
Best Local Similarity 100.0%; Pred.No. 3.1e-29;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1507 CAAGCTTGCTTTTGGCGGATGAGAAATTTTCAGCTGATACAGATTAAATCAGA 1566
    |||
Db 436 CAAGCTTGCTTTTGGCGGATGAGAAATTTTCAGCTGATACAGATTAAATCAGA 495
    |||
QY 1567 ACGCAGAACGGCTCTGATTAACAGAAATTTGCTGGCGGAGTAGCGCGGTGTCACC 1626
    |||
Db 496 ACGCAGAACGGCTCTGATTAACAGAAATTTGCTGGCGGAGTAGCGCGGTGTCACC 555
    |||
QY 1627 TGACCCCATGCCGAATCTCAGAAAGTGAA 1653
    |||
Db 556 TGACCCCATGCCGAATCTCAGAAAGTGAA 582
    |||

```

```

RESULT 15
US-09-131-028A-6/c
; Sequence 6, Application US/09131028A
; Patent No. 6287866
; GENERAL INFORMATION:
;   APPLICANT: Abbott Laboratories
;   APPLICANT: Mukerji, Pradip
;   APPLICANT: Lemmel, Steven A.
;   APPLICANT: Leonard, Amanda Eun-Yeong
;   APPLICANT: Chaudhary, Sunita
;   TITLE OF INVENTION: BETA-CASEIN EXPRESSING CONSTRUCTS
;   FILE REFERENCE: 6004.US.PI
;   CURRENT APPLICATION NUMBER: US/09/131,028A
;   CURRENT FILING DATE: 1998-08-07
;   PRIOR APPLICATION NUMBER: US 08/064,440
;   PRIOR FILING DATE: 1993-05-21
;   NUMBER OF SEQ ID NOS: 22
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 6
;   LENGTH: 8430
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (6124)...(6939)
;   OTHER INFORMATION: KANR - kanamycin resistance in forward orientation
US-09-131-028A-6

```

```

Query Match      8.9%; Score 147; DB 4; Length 8430;
Best Local Similarity 100.0%; Pred.No. 4.1e-29;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1507 CAAGCTTGCTTTTGGCGGATGAGAAATTTTCAGCTGATACAGATTAAATCAGA 1566
    |||
Db 7428 CAAGCTTGCTTTTGGCGGATGAGAAATTTTCAGCTGATACAGATTAAATCAGA 7369
    |||
QY 1567 ACGCAGAACGGCTCTGATTAACAGAAATTTGCTGGCGGAGTAGCGCGGTGTCACC 1626
    |||
Db 7368 ACGCAGAACGGCTCTGATTAACAGAAATTTGCTGGCGGAGTAGCGCGGTGTCACC 7309
    |||
QY 1627 TGACCCCATGCCGAATCTCAGAAAGTGAA 1653
    |||
Db 7308 TGACCCCATGCCGAATCTCAGAAAGTGAA 7282
    |||

```

Search completed: January 16, 2003, 08:53:08
Job time : 88 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 08:18:51 ; Search time 70 Seconds
(without alignments)
10528.861 Million cell updates/sec

Title: US-09-830-111A-1
Perfect score: 1653
Sequence: 1 ttttggggtcgaaagtc.....atgccgaactcagaagtga 1653

Scoring table: IDENTITY NUC

Gapop 10.0 , Capext 1.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMP.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMP.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMP.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMP.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMP.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMP.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMP.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	179.2	10.8	5926	9	US-10-024-809-3
C 2	179.2	10.8	5926	9	US-10-024-809-4
C 3	145.4	8.8	6271	10	US-09-993-292A-1
C 4	145.4	8.8	8908	10	US-09-993-292A-18
C 5	138	8.3	425	10	US-09-960-428-10
C 6	136	8.2	1173	10	US-09-887-576-678
C 7	136	8.2	1179	9	US-09-938-842A-990
C 8	130.8	7.9	7416	10	US-09-900-575-46
C 9	120.8	7.3	894	9	US-09-938-842A-1124
C 10	115.4	7.0	799	10	US-09-912-020-27
C 11	79.2	4.8	927	10	US-09-974-300-999
C 12	68.2	4.1	972	10	US-09-815-242-9636
C 13	64.2	3.9	787	12	US-10-040-037-1
C 14	63	3.8	1323	12	US-10-013-036-3
C 15	61.6	3.7	103	10	US-09-179-536B-267
C 16	61.6	3.7	513	10	US-09-974-300-5444
C 17	61.6	3.7	1361	10	US-09-393-634-14
C 18	61.4	3.7	293	9	US-10-132-561-11
C 19	61.4	3.7	293	9	US-10-132-561-12

20	61.4	3.7	458	9	US-10-132-561-14	Sequence 14, Appl
C 21	61.4	3.7	458	9	US-10-132-561-14	Sequence 14, Appl
C 22	61.4	3.7	459	9	US-10-132-561-13	Sequence 13, Appl
C 23	61.4	3.7	600	9	US-10-132-561-5	Sequence 5, Appl
C 24	61.4	3.7	789	9	US-10-132-561-16	Sequence 16, Appl
C 25	61.4	3.7	789	9	US-10-132-561-16	Sequence 16, Appl
C 26	61.4	3.7	790	9	US-10-132-561-15	Sequence 15, Appl
C 27	61.4	3.7	790	9	US-10-132-561-15	Sequence 15, Appl
C 28	61.4	3.7	2686	10	US-09-785-269-7	Sequence 7, Appl
C 29	61	3.7	71	10	US-09-380-932-3	Sequence 3, Appl
C 30	61	3.7	206	9	US-10-033-297-27	Sequence 27, Appl
C 31	61	3.7	206	9	US-10-081-806-32	Sequence 32, Appl
C 32	61	3.7	586	10	US-09-764-847-1873	Sequence 1873, Ap
C 33	61	3.7	972	10	US-09-815-242-6224	Sequence 6224, Ap
C 34	61	3.7	6741	10	US-09-794-384A-14	Sequence 14, Appl
C 35	60.8	3.7	560	10	US-09-772-134B-26	Sequence 26, Appl
C 36	60	3.6	75	10	US-09-955-649-6	Sequence 6, Appl
C 37	60	3.6	75	10	US-09-973-013-6	Sequence 6, Appl
C 38	60	3.6	119	8	US-08-852-020-15	Sequence 15, Appl
C 39	60	3.6	149	12	US-10-012-211A-5	Sequence 5, Appl
C 40	60	3.6	298	9	US-09-808-124A-4	Sequence 4, Appl
C 41	60	3.6	298	10	US-09-845-157-6	Sequence 6, Appl
C 42	60	3.6	384	10	US-09-757-992-6	Sequence 6, Appl
C 43	60	3.6	384	10	US-09-757-992-13	Sequence 13, Appl
C 44	60	3.6	910	8	US-08-852-020-5	Sequence 5, Appl
C 45	59.6	3.6	102	9	US-10-015-219-106	Sequence 106, App

ALIGNMENTS

RESULT 1
US-10-024-809-3/c
; Sequence 3, Application US/10024809
; Publication No. US20020188105A1
; GENERAL INFORMATION:
; APPLICANT: CRAIG, NANCY L
; TITLE OF INVENTION: GAIN OF FUNCTION MUTATIONS IN
; ATP-DEPENDENT TRANSDUCTION PROTEINS

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Anne Brown (Alston & Bird, LLP)

STREET: 3605 Glenwood Ave.

CITY: Raleigh

STATE: NC

COUNTRY: USA

ZIP: 27608

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/10/024,809

FILING DATE: 19-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/027,169

FILING DATE: 1998-FEB-20

ATTORNEY/AGENT INFORMATION:

NAME: Brown, Anne

REGISTRATION NUMBER: 36,463

REFERENCE/DOCKET NUMBER: 5789-3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919 420 2205

TELEFAX: 919 881 3175

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 5926 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: circular

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "pem delta R.adj to 1"
SEQUENCE DESCRIPTION: SEQ ID NO: 3
US-10-024-809-3

Query Match 10.8%; Score 179.2; DB 9; Length 5926;
Best Local Similarity 91.3%; Pred. No. 6.5e-45;
Matches 190; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1446 GGTCAAGATAGGAATTGAGCTCGGTACCCGGGATCCTTGAAGTCGACCTGCAGGCAT 1505
DB 4282 GTTATTTAAAGCTGTCCGACGACCTTTCTTGGGATCCTCTAGAGTCGACCTGCAGGCAT 4223
QY 1506 GCAAGCTTGGCTGTTTGGCCGATGAGAGAATTTTGAAGCTTGAATTAATCAG 1565
DB 4222 GCAAGCTTGGCTGTTTGGCCGATGAGAGAATTTTGAAGCTTGAATTAATCAG 4163
QY 1566 AACGAGAAGCGGTCTGATTAACAAGATTTTGCCTGGCGGAGTAGCGCGGTGTCAC 1625
DB 4162 AACGAGAAGCGGTCTGATTAACAAGATTTTGCCTGGCGGAGTAGCGCGGTGTCAC 4103
QY 1626 CTGACCCCATGCCGAAGTGAAGTGA 1653
DB 4102 CTGACCCCATGCCGAAGTGAAGTGA 4075

RESULT 2

US-10-024-809-4/c
Sequence 4, Application US/10024809
Publication No. US20020188105A1

GENERAL INFORMATION:

APPLICANT: CRAIG, NANCY L
TITLE OF INVENTION: GAIN OF FUNCTION MUTATIONS IN
ATP-DEPENDENT TRANSPORT PROTEINS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Anne Brown (Alston & Bird, LLP)
STREET: 3605 Glenwood Ave.
CITY: Raleigh
STATE: NC
COUNTRY: USA
ZIP: 27608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/024,809

FILING DATE: 19-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/027,169

FILING DATE: 1998-FEB-20

ATTORNEY/AGENT INFORMATION:
NAME: Brown, Anne

REGISTRATION NUMBER: 36,463

REFERENCE/DOCKET NUMBER: 5789-3

TELECOMMUNICATION INFORMATION:
TELEPHONE: 919 420 2205

TELEFAX: 919 881 3175

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 5926 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: circular

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "pem-delta"

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-10-024-809-4

Query Match 10.8%; Score 179.2; DB 9; Length 5926;
Best Local Similarity 91.3%; Pred. No. 6.5e-45;

Matches 190; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1446 GGTCAAGATAGGAATTGAGCTCGGTACCCGGGATCCTTGAAGTCGACCTGCAGGCAT 1505
DB 3918 GTTATTTAAAGCTGTCCGACGACCTTTCTTGGGATCCTCTAGAGTCGACCTGCAGGCAT 3859
QY 1506 GCAAGCTTGGCTGTTTGGCCGATGAGAGAATTTTGAAGCTTGAATTAATCAG 1565
DB 3858 GCAAGCTTGGCTGTTTGGCCGATGAGAGAATTTTGAAGCTTGAATTAATCAG 3799
QY 1566 AACGAGAAGCGGTCTGATTAACAAGATTTTGCCTGGCGGAGTAGCGCGGTGTCAC 1625
DB 3798 AACGAGAAGCGGTCTGATTAACAAGATTTTGCCTGGCGGAGTAGCGCGGTGTCAC 3739
QY 1626 CTGACCCCATGCCGAAGTGAAGTGA 1653
DB 3738 CTGACCCCATGCCGAAGTGAAGTGA 3711

RESULT 3

US-09-993-292A-1/c
Sequence 1, Application US/09993292A
Patent No. US20020146430A1

GENERAL INFORMATION:

APPLICANT: James E. Galen
TITLE OF INVENTION: USE OF CLY A HEMOLYSIN FOR EXCRETION OF
PROTEINS
FILE REFERENCE: UOFMD.007A
CURRENT APPLICATION NUMBER: US/09/993,292A
CURRENT FILING DATE: 2001-11-23
PRIOR APPLICATION NUMBER: 60/252,516
PRIOR FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 6271
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: pSEC84 Expression Plasmid

US-09-993-292A-1

Query Match 8.8%; Score 145.4; DB 10; Length 6271;
Best Local Similarity 99.3%; Pred. No. 2e-34;
Matches 146; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1507 CAAGCTTGGCTGTTTGGCCGATGAGAGAATTTTCAAGCTTGAATTAATCAGA 1566
DB 2571 CCAGCTTGGCTGTTTGGCCGATGAGAGAATTTTCAAGCTTGAATTAATCAGA 2512
QY 1567 ACGAGAAGCGGTCTGATTAACAAGATTTTGCCTGGCGGAGTAGCGCGGTGTCAC 1626
DB 2511 ACGAGAAGCGGTCTGATTAACAAGATTTTGCCTGGCGGAGTAGCGCGGTGTCAC 2452
QY 1627 TGACCCCATGCCGAAGTGAAGTGA 1653
DB 2451 TGACCCCATGCCGAAGTGAAGTGA 2425

RESULT 4

US-09-993-292A-18/c
Sequence 18, Application US/09993292A
Patent No. US20020146430A1

GENERAL INFORMATION:

APPLICANT: James E. Galen
TITLE OF INVENTION: USE OF CLY A HEMOLYSIN FOR EXCRETION OF
PROTEINS
FILE REFERENCE: UOFMD.007A
CURRENT APPLICATION NUMBER: US/09/993,292A
CURRENT FILING DATE: 2001-11-23
PRIOR APPLICATION NUMBER: 60/252,516

; PRIOR FILING DATE: 2000-11-22
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 18
 ; LENGTH: 8908
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: pSEC84sacB vector
 US-09-993-292A-18

Query Match 8.8%; Score 145.4; DB 10; Length 8908;
 Best Local Similarity 99.3%; Pred. No. 2.5e-34;
 Matches 146; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1507 CAAGCTTGGCTGTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGA 1566
 Db 5208 CCAGCTTGGCTGTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGA 5149
 QY 1567 ACCGAGAGCGGCTGATATAACAGATTTCCTGGCGGAGTAGCGCGGTGGTCCCAACC 1626
 Db 5148 ACGCAGAGCGGCTGATATAACAGATTTCCTGGCGGAGTAGCGCGGTGGTCCCAACC 5089
 QY 1627 TGACCCCATCGGAACTCAGAAGTGAA 1653
 Db 5088 TGACCCCATCGGAACTCAGAAGTGAA 5062

RESULT 5
 US-09-960-428-10
 ; Sequence 10, Application US/09960428
 ; Patent No. US20020115147A1

; GENERAL INFORMATION:
 ; APPLICANT: Roche Diagnostics GmbH
 ; TITLE OF INVENTION: Method for producing an active heterodimeric AMV-RT in prokaryotic
 ; CURRENT APPLICATION NUMBER: US/09/960,428
 ; CURRENT FILING DATE: 2001-09-21
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 10
 ; LENGTH: 425
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 US-09-960-428-10

Query Match 8.3%; Score 138; DB 10; Length 425;
 Best Local Similarity 100.0%; Pred. No. 8.1e-33;
 Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1516 CTGTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAAGCGCAAG 1575
 Db 1 CTGTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAAGCGCAAG 60
 QY 1576 CGGCTGATATAACAGAAATTTGCTGCGGAGTAGCGCGGTGGTCCACCTGACCCCAT 1635
 Db 61 CGGCTGATATAACAGAAATTTGCTGCGGAGTAGCGCGGTGGTCCACCTGACCCCAT 120
 QY 1636 GCCGAACCTCAGAAGTGAA 1653
 Db 121 GCCGAACCTCAGAAGTGAA 138

RESULT 6
 US-09-887-576-678
 ; Sequence 678, Application US/09887576
 ; Patent No. US20020144047A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Budworth, P.
 ; APPLICANT: Brown, D.
 ; APPLICANT: Chang, H.
 ; APPLICANT: Zhu, T.
 ; APPLICANT: Han, B.

; APPLICANT: Wang, X.
 ; APPLICANT: Cooper, Bret
 ; TITLE OF INVENTION: Promoters for regulation of plant expression
 ; FILE REFERENCE: 1360.001US1
 ; CURRENT APPLICATION NUMBER: US/09/887,576
 ; CURRENT FILING DATE: 2001-06-25
 ; PRIOR APPLICATION NUMBER: US 60/213,848
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: US 60/214,087
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: US 60/258,692
 ; PRIOR FILING DATE: 2000-12-29
 ; NUMBER OF SEQ ID NOS: 875
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 678
 ; LENGTH: 1173
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-887-576-678

Query Match 8.2%; Score 136; DB 10; Length 1173;
 Best Local Similarity 50.3%; Pred. No. 6.2e-32;
 Matches 380; Conservative 0; Mismatches 355; Indels 21; Gaps 1;

QY 671 CAACGGCGGTTGGCTGAGATCAGGAGATGATCCATGCAGCATCACTCTCCAGCAGCAC 730
 Db 409 CACCGCGGTTGGCTGAGATCAGGAGATGATCCATGCAGCATCACTCTCCAGCAGCAC 468
 QY 731 GTTATCGAGCTTCCGAGACCGACGAAACGCACCATCCGAAACAGGATTCGGAAC 790
 Db 469 GTGTTAGCAGAGAGTGCATGCGAAGAGAAACAGTTTCATGAGCTTTTCGGCACA 528
 QY 791 AAGATGCGGATTTTGGCTGGTGATTTCTTGTGGGAGCGGCTCTGTTGATTTGGCGAGG 850
 Db 529 AGAGTAGCGGTGCTAGCTGGAGATTTTCATGTTGCTCAAGCGTCTGCTAGTCAAT 588
 QY 851 TTGCGCAATCCGAGGTGATGAGCTTTTGGCTACTGTTTATGCAAACTTGGTTGAGGGA 910
 Db 589 CTCGAGAACTTGAAGTTATTAAAGTCATCAGTCAGGTGATCAAGAGACTTTGCAAGCGGA 648
 QY 911 GAGTTCATGCAGTTGAAATAATCTGTTGATGATGCGATTGAGGCTACGCGGACGAGAA 970
 Db 649 GAGATAAGCAGCGGCTCAGCTTATTGCTGCG-----ACACC 687
 QY 971 ACCTTCGATTACTATTTCGAGAAAGACTTACTTGAAGACTCGCTCTTGTGTTGCAAGTCG 1030
 Db 688 AAGCTCGACGAGTACTTACTCAAAAGTTTCTACAAGACACCTCTTTAGTGGCTGCGAGC 747
 QY 1031 TGAGAGCAAGTCGCTTCTGGGTGGTGCTACGCTGAGGTGCTGATGCTGCTTATGCT 1090
 Db 748 ACCAAAGGAGCTGCCATTTTTCAGCAGAGTTGAGCCTGATGTGACAGAAATGTACGAG 807
 QY 1091 TAGCGAAGGAACCTTGGTTGGCATTTCCAGATCGTCAGCAGATGCTCGACTACCGTC 1150
 Db 808 TTTGGGAAGAATCTCGGTCTCTTTCCAGATAGTTGATGATATTTTGGATTTCACCTCAG 867
 QY 1151 TCCGCTACCGACCTCGGTAAAGCCCGGTCAGACCTCCAGCTCGCTCTCGCCACCGCG 1210
 Db 868 TCAGCAGACGAGCTCGGGAAGCCAGCAGGAGTGATTTGGCTAAAGTAACTTAAACAGCA 927
 QY 1211 CCGGCCCTCTTTCGATGGAAGACACCGAGCTCGTCCCATGATCAAGCGCAAGTTC 1270
 Db 928 CCTGTGATTTTCTGCTCTGGAGAGGAGCCAAAGGCTTAAGAGAGATCATTTGAGTCAGAGTTT 987
 QY 1271 TCTGACCCAGGAGACGCTCGAGCGGTGACCGAGTTGGTCGAGAAAGTATGATTTGGAG 1330
 Db 988 TGTGAGCGCGGTTCTCTGGAAGAAAGCGATTGAAGCGGTGCAAAAGGTTGGGGGATTAAG 1047
 QY 1331 AAGACGAGAGCCCTTGGCGGAGGAGTATGCCAGAAAGCGGTTGGATGCAATTCGAGCGTTC 1390
 Db 1048 AGAGCACAAGAATTGGCTAGGAGAAAGCTGATGACGCTATAAAGATCTACAGTGTCTA 1107
 QY 1391 CCGGAGAGTCCGGCAGCGAAGGCTTTGGAGCAGTTG 1426

Db 1108 CCTCGAAGTGGCTTCAGGTGGCTCTAGAAATATG 1143

RESULT 7

US-09-938-842A-990
 ; Sequence 990, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kieps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRIPT300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 990
 ; LENGTH: 1179
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-938-842A-990

Query Match 8.2%; Score 136; DB 9; Length 1179;
 Best Local Similarity 50.3%; Pred. No. 6.2e-32;
 Matches 380; Conservative 0; Mismatches 355; Indels 21; Gaps 1;

Qy 671 CAACGCGGCTGGCTGAGATCAGGAGATGATCCATGACATCTCTCCACGACGAC 730
 Db 415 CACCGCGCTTGGCTGAGATCAGGAGATGATCCATGACATCTCTCCACGACGAT 474
 Qy 731 GTTATGACGCTTCGAGACGACGAAACGACCATCCGGAACCGAGCATTCGGAAC 790
 Db 475 GTTATGACGAGATGATCAGGAGAGAAAGAAACGATTCATGAGCTTTTCGACACA 534
 Qy 791 AAGATGCGCATTTTGGCTGATGATTTCTTTGGACGCGGCGCTCTGTCATTGGCGAG 850
 Db 535 AGAGTAGCGGTCTGAGTGAATTTCTTTGCTCAAGCGCATGTAAGTATGACAAAT 594
 Qy 851 TTGCGCAATCCGAGAGTATGAGCTTTGGCTACTGTTATTCGAACTTTGGTTAGGGA 910
 Db 595 CTCGGAATCTTGAAGTATTAAGCTCATGATCAAGTATCAAAAGCTTTGCAAGCGGA 654
 Qy 911 GAGTTCATGCACTTGAAGAAATACTGTTGATGATGAGATTGAGGCTACGCGACGAGAA 970
 Db 655 GAGATTAAGCAGGCTCCAGCTTATTTGACTGCG-----ACACC 693
 Qy 971 ACGTTGATTACTATTGCGAGAGACTTACTTAAGACTGCGCTCTTGAATTGCCAAGTGC 1030
 Db 694 AAGCTCGAGAGTACTTACTTCAAAAGTTCTTAAGACAGCGCTTTAGTGGCTGAGAG 753
 Qy 1031 TGCAGAGCAAGTCCGCTTGGGTGCTACGCTGAGGTTGCTGATCTGCTTATGCT 1090
 Db 754 ACCAAGAGGCTGCCATTTTTCAGCAGAGTTGAGCCTGATGTGACAGAACAAATGTACGAG 813
 Qy 1091 TACGGAAGAACCTTGGTTGGATTCCGATCGTGCAGCAATGCTGCATACACCGTC 1150
 Db 814 TTGGGAAGAAATCTGCTCTCTTTCCAGATAGTTGATGATTTTGGATTTCACCTGAG 873
 Qy 1151 TCCGCTACCGACCTCGGTAAAGCCCGCGGTGACAGCTTCAGCTCGCTCTCCGACCGCG 1210
 Db 874 TCGAAGAGCACTCGGGAAGCCAGAGGAGTGAATTTGGCTAAAGTAACTTAACAGA 933
 Qy 1211 CCGGCGCTCTTGGCATGGAAGACCAAGCGGAGCTCGGTCCCATGATCAAGCGCAAGTTTC 1270

Db 934 CCTGATTTTTCGCTCTGAGAGGAGGCAAGGCTTAAGAGATCATTTGATCAGATT 993

Qy 1271 TCTGACCCGAGAGAGCTGACGCTGACGCGAGTTGTTGAGAAAAGTATGATTTGAG 1330

Db 994 TGTGAGCGGCTTCTCTGGAAGAACGATTAAGCGGTGACAAAAGGTGGGGGATTAAAG 1053

Qy 1331 AAGAGCAAGGCTTGGCGGAGAGATGCCAGAAAGCGTTGATGATTCGACGCTTC 1390

Db 1054 AGAGCAAGAAATTTGGCTTGAAGGAGAAAGCTGATGACGCTTAAAGAAATCTACAGTCTCA 1113

Qy 1391 CCGGAGAGTCCGCAAGGAGGAGGCTTTGGAGCAGTTG 1426

Db 1114 CCTGAAGTGGCTTCAGGTGGCTCTTAGAAGATATG 1149

RESULT 8

US-09-900-575-46
 ; Sequence 46, Application US/09900575
 ; Patent No. US20020150587A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Langermann, Solomon
 ; APPLICANT: Revel, Andrew
 ; APPLICANT: Auguste, Christine
 ; APPLICANT: Burteln, Jeanne
 ; TITLE OF INVENTION: Film Adhesin Proteins and Methods of Use
 ; FILE REFERENCE: 469201-549
 ; CURRENT APPLICATION NUMBER: US/09/900,575
 ; CURRENT FILING DATE: 2001-07-06
 ; PRIOR APPLICATION NUMBER: US/60/216,750
 ; PRIOR FILING DATE: 2000-07-07
 ; NUMBER OF SEQ ID NOS: 64
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 46
 ; LENGTH: 7416
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Sequence of Plasmid pCGA126-1
 ; US-09-900-575-46

Query Match 7.9%; Score 130.8; DB 10; Length 7416;
 Best Local Similarity 98.5%; Pred. No. 7.5e-30;
 Matches 132; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1512 TTGCGTCTTTGGCGGATGAGAAAGATTTCAAGCTGATACAGATTAAATCAGAACGCA 1571

Db 2255 TTCCCTGTTTGGCGGAGAGAGAAAGATTTCAAGCTGATACAGATTAAATCAGAACGCA 2314

Qy 1572 GAAGCGTCTGATAAACAGAAATTTGCTGGCGGAGAGCGGCTGCTCCACCTGACC 1631

Db 2315 GAAGCGTCTGATAAACAGAAATTTGCTGGCGGAGAGCGGCTGCTCCACCTGACC 2374

Qy 1632 CCATGCCGAAGTCA 1645

Db 2375 CCATGCCGAAGTCA 2388

RESULT 9

US-09-938-842A-1124
 ; Sequence 1124, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kieps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRIPT300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24

```

; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1124
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-1124

Query Match
Best Local Similarity 7.3%; Score 120.8; DB 9; Length 894;
Matches 242; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

QY 954 CTACGGCAGCGAGAAAGCTTCGATTACTTATTTGAGAGACTTACTTTGAAGACTGCGT 1013
DB 374 CAACCGAGCAGCGTTATGATGACTACTTACATGAGAGACATATTATAAGACAGCAT 433

QY 1014 CTTGATTGCCAAGTCGTGAGAGCAAGTCGCTTCTGGGTGTGCTACGCTGAGGTTG 1073
DB 434 CGCTAATCTTAACAGCTGCAAAAGCTGTTGCCGTCTCTCACTGGACAAACAGCAGAAAGTTG 493

QY 1074 CTGATCTGCTTATGCTTACGGAAGAACTTGGTTTGGCATTCAGATCGTCGACGACA 1133
DB 494 CCGTGTAGCTTTTGTAGTGGAGGAATCTGGGTTTAGCAITTCCTTAATAGACGACA 553

QY 1134 TGCTGATACACGCTTCGCTTACCGACCTCGGTAAAGCCGCGGTGACAGCTCCAGC 1193
DB 554 TTCTTGATTTCACGGCAGCATCTGCTCTCTCGGAAGGATGCTTGTGATGATTCGCC 613

QY 1194 TCGGTCTCGCACCGCGGCGCTCTTCGATGGAAGACACACGCGGAGCTCGGTCCCA 1253
DB 614 ATGGAGTCAATACAGCGCCCAATCTCTTTGCCATGGAAGAGTTTCTCAACTACGCGAAG 673

QY 1254 TGATCAAGCGCAAGTTCTCTGACCCAGGAGACGTCGAGCGTCGACGCGAGTTGGTCGAGA 1313
DB 674 TTGTTGATCAAGTTGAAAAGATCTTAGGAATGTTGACATGCTTTAGAGATATCTTTGGGA 733

QY 1314 AAAGTGATGGATTGGAGAGACGAGAGCCTTGGCGGAGGAGTATGCCAGAGGCGGTTGG 1373
DB 734 AGAGCAAGGGAATACAGAGGCGCAAGAGATAGCCATGGAACATGCAATCTAGACAGCAG 793

QY 1374 ATGCAATTCGAGCTTCCCGGAGA 1397
DB 794 CTGCAATCGGGTCTTACTCTGAAA 817

RESULT 10
US-09-912-020-27
; Sequence 27, Application US/09912020
; Patent No. US20020045592A1
; GENERAL INFORMATION:
; APPLICANT: Zyskind, Judith
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Trawick, John
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Froelich, Jamie M.
; APPLICANT: Cart, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; FILE REFERENCE: ESCHERICHIA COLI
; CURRENT APPLICATION NUMBER: US/09/912,020
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: 09/492,709
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/117,405
; PRIOR FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27

; LENGTH: 799
; TYPE: DNA
; ORGANISM: E. Coli
; US-09-912-020-27

Query Match
Best Local Similarity 7.0%; Score 115.4; DB 10; Length 799;
Matches 116; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1537 GATTTTCAGCCTGATACAGATTAAATTCAGAACGCGAAGCGGTCTCTGATAAACAGAAATTT 1596
DB 475 GCTTTTCAGCCTGATACAGATTAAATTCAGAACGCGAAGCGGTCTCTGATAAACAGAAATTT 534

QY 1597 GCTGCGCAGTAGCGCGGTGTCCTCCACTGACCCCATGCCGAACTCAGAAAGTGAA 1653
DB 535 GCTGCGCAGTAGCGCGGTGTCCTCCACTGACCCCATGCCGAACTCAGAAAGTGAA 591

RESULT 11
US-09-974-300-999
; Sequence 999, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: Expression
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 999
; LENGTH: 927
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
; US-09-974-300-999

Query Match
Best Local Similarity 4.8%; Score 79.2; DB 10; Length 927;
Matches 341; Conservative 0; Mismatches 363; Indels 24; Gaps 2;

QY 694 GGAGATGATCCATGCGAGCATCTCTCCACGACGAGTTCATCGCGCTTCCGAGACCAG 753
DB 221 GGAATTGATACATATGGCTTCGTCGTACACGACGATGTCATCGATGCGGAGCTGAG 280

QY 754 AGCAACGCAACCATCCGAAACAGGCATTCGGAACCAAGATGCGGATTTTGGCTGGTGA 813
DB 281 AGAGGACAGCGACATCAAGCGAAGTGGATTAACCGGATTCGAATGTATACAGGTGA 340

QY 814 TTTCTTTTGGGACGCGGCGTCTGTTGATGCGAGGTTTCCGCAATCCGAGGTGATGA 873
DB 341 TTATTGTTGCGTCTGCTCACTTGAAGCGGATGACAAAAATAAACGAAACCCGATGA 400

QY 874 GCTTTTGGCTACTGTTTATTCGAACTTGTGTTGAGGAGAGTTCATCGAGTTGAAAAATAC 933
DB 401 AATCTTGTCCAAGGCGATTGTGAAAGTCTGCTTTGGAGAAATTTGAACAAATTAAGATAA 460

QY 934 TGTGATGATGCGATTGAGGCTACGCGCAGCAGGAAACGTTTCGATTACTATTTCAGAA 993
DB 461 ATATAAT-----ATGAGCAAAATCTCAGAACCTATTTTGAGAAG 499

QY 994 GACTTACTTGAAGACTCGCTCTTGAATGCCAAGTCTGTCAGAGCAAGTCCGCTTCTGGG 1053
DB 500 AATCAGGCGCAAGACAGCACTTCTAATCGCTGAGCTGCCAGCTTGGAGCCATTCCTGC 559

QY 1054 TGGTGTACGCTGAGGTTGCTGATGCTCTTATGCTTACGGAAGCACTTGGTTTGGC 1113
DB 560 TGGTGGGCTGAAACCAITTCATAAAACGTTGTACTGCTTTGGGTATTATGTGCGCATGTC 619

```

QY 1114 ATTCAGATCGTCAGCAGCATCTCGACTACACCGTCTCCGCTACCGACCTGGTAGCC 1173
 DB 620 CTATCAATTTATTTGATGATCTCTGATTTCAATCAAGCAAGCAAGAGCTTGGCAAGCA 679
 QY 1174 CGCCGTCGACGCTCCGAGCTGGTCTGCCACCGCGCCGCTCTTGGCATGGAAGCA 1233
 DB 680 TGTGGGACCGCATCTGCTTCAGGGGAATGTCAAGCTTCCGATTTATGTCAAGGAATC 739
 QY 1224 CCACGCCGAGCTCGGCTCCCATGATCAAGCGCAA---GTTCTCTGACCCAGAGAGCTCGA 1290
 DB 740 GCTTAGTAAAAAGCGCGCTGATGTGTCAACAGTAAACACCTCCAGAAACAGATCA 799
 QY 1291 GCGTCAACGCGAGTGTGTCGAGAAAAAGTAGATGGAAGAGAGAGAGCTTGGCGGA 1350
 DB 800 GCGAGTGAATCGAAGAGCTGAGAAAAAGAGATGCAATTTGAACAATCGTTCAGGGGTCACTGA 859
 QY 1351 GAGTATGCCCCGAGAGGCTTGGATGCAATTTGGACGTTCCCGAGAGATCCCGACCGGA 1410
 DB 860 AATGATTTTAAACAAGGCTTTCGAATTTGGTAAACGAACCGATTAACCGGAGCGCCAC 919
 QY 1411 GCGTTGG 1418
 DB 920 CCGAGTGG 927

RESULT 12

US-09-815-242-9636
 / Sequence 9636, Application US/09815242
 / Patent No. US2002061569A1
 / GENERAL INFORMATION:
 / APPLICANT: Haselbeck, Robert
 / APPLICANT: Ohlsen, Karl L.
 / APPLICANT: Zyskind, Judith W.
 / APPLICANT: Wall, Daniel
 / APPLICANT: Trawick, John D.
 / APPLICANT: Carr, Grant J.
 / APPLICANT: Yamamoto, Robert T.
 / APPLICANT: Xu, H. Howard
 / TITLE OF INVENTION: Identification of Essential Genes in
 / TITLE OF INVENTION: Prokaryotes
 / FILE REFERENCE: ELITRA.011A
 / CURRENT APPLICATION NUMBER: US/09/815,242
 / PRIOR FILING DATE: 2001-03-21
 / PRIOR APPLICATION NUMBER: 60/191,078
 / PRIOR FILING DATE: 2000-03-21
 / PRIOR APPLICATION NUMBER: 60/206,848
 / PRIOR FILING DATE: 2000-05-23
 / PRIOR APPLICATION NUMBER: 60/207,727
 / PRIOR FILING DATE: 2000-05-26
 / PRIOR APPLICATION NUMBER: 60/242,578
 / PRIOR FILING DATE: 2000-10-23
 / PRIOR APPLICATION NUMBER: 60/253,625
 / PRIOR FILING DATE: 2000-11-27
 / PRIOR APPLICATION NUMBER: 60/257,931
 / PRIOR FILING DATE: 2000-12-22
 / PRIOR APPLICATION NUMBER: 60/269,308
 / PRIOR FILING DATE: 2001-02-16
 / NUMBER OF SEQ ID NOS: 14110
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 9636
 / LENGTH: 972
 / TYPE: DNA
 / ORGANISM: Salmonella typhi
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: (1)...(972)
 / US-09-815-242-9636

Query Match 4.1%; Score 68.2; DB 10; Length 972;
 Best Local Similarity 45.8%; Pred. No. 5.9e-11;
 Matches 348; Conservative 0; Mismatches 378; Indels 33; Gaps 2;

QY 667 CTCGCAACGGCGGTTGGCTGATCAAGAGATGATTCATGACATCACTCTCCACGA 726
 DB 192 CGCGACGTCGCTATCGCGCTCTGATGAGTTATTCACACGCGACGCTGCTACATGA 251
 QY 727 CGAGCTTATCGACCTTCCGAGACGACGAAACGACCATCCGAAAACGAGGATTCGG 786
 DB 252 TGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 311
 QY 787 AAACAAGATGGCATTTTGGCTGATGATTTCTTGTGGAGCGGCGTCTGTGATTCGC 846
 DB 312 TAAATGCCCGCCAGCGTGTGTGGGAGCTTATATACCCGCGCTTCCAGATGATGAC 371
 QY 847 GAGGTTGGCAATCCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 906
 DB 372 CAGCTCGGCTGCTTAAAGTGTGAGATGATGATGATGATGATGATGATGATGATGATG 431
 QY 907 GGGAGATTCAATGCAATGTAATAAATACCTGTGATGATGATGATGATGATGATGATG 966
 DB 432 AGCGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 474
 QY 967 GGAAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1026
 DB 475 ---ACGAAAGAAATATCATGCGGCTGATTTACGTAACGCGCGCTCTGTTCGAAGC 530
 QY 1027 GTCTGCAAGACGATGAGCTTCTGGGATGCTACGCTGAGGTTGCTGATGCTGTTA 1086
 DB 531 CGCGCCCGAGTGTCTCGGTATTTCTTCCGTTGTACGCTTGACGAAGAAAGGTTGCA 590
 QY 1087 TGCTTACGGAAGAAACCTTGTGTTGGATTCAGATTCGTGACGATGCTGACTGAC 1146
 DB 591 GAGCTATGCGCTATGCTTGTGATGAGCTTTCAGCTCATTTGACATGCTGCTGATTA 650
 QY 1147 GCTTCTCCGATACCACTCTGATGAGCCCGGCTGCAAGACTTCCAGCTGCTGCTG 1206
 DB 651 TGCCACGCGCGACATCTCGTAAATAATGTGGTATGATCCATGATGAGGCPAACTAC 710
 QY 1207 CGCGCGCGGCTCTTGGATGAGAACACACGCGGAGCTCGGCTCCATGATCAACG 1266
 DB 711 CTTACCGTGTCTTCAAGGCAATGCGGACGATGACGCAAGAACTGATGATGATGATG 770
 QY 1267 GTTCTCTGACCCGAGAGAGCTGACGCTGACGCGAGTGTGCGAATAA----- 1316
 DB 771 CGCTATGAAACAAGGTAATGCGGCTCATCTCTGAAACCGGTTCTGGAAGCATGAC 830
 QY 1317 -GTGATGATGAGAAAGACGAGACCTTGGCGGAGATGATGCTCCAGAGGCTTGA 1374
 DB 831 CTGCGGCTGCTGGAATGAGCGCTCAGGAGCGGAAAGAACCGCAAGCGATATC 890
 QY 1375 TGCATTTGCGAGCTTCCGGAAGATCCGSCACGGAAGGC 1413
 DB 891 CCGGTTGCAGATATTGCGGATACCCCATGCGTGAAGC 929

RESULT 13

US-10-040-037-1
 / Sequence 1, Application US/10040037
 / Patent No. US20020102593A1
 / GENERAL INFORMATION:
 / APPLICANT: Leonard, Jack T.
 / APPLICANT: MacDonald, Constance
 / APPLICANT: Gabriels, Joseph
 / TITLE OF INVENTION: Method for Sequencing Reaction Cleanup
 / TITLE OF INVENTION: By Constant Pressure Differential Ultrafiltration
 / FILE REFERENCE: MCA-480A
 / CURRENT APPLICATION NUMBER: US/10/040,037
 / PRIOR FILING DATE: 2001-10-19
 / PRIOR APPLICATION NUMBER: 60/154,448
 / PRIOR FILING DATE: 1999-09-17
 / NUMBER OF SEQ ID NOS: 1
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 1
 / LENGTH: 767
 / TYPE: DNA

```

; ORGANISM: Escherichia coli bacteriophage lambda
US-10-040-037-1
Query Match          3.9%; Score 64.2; DB 12; Length 767;
Best Local Similarity 95.7%; Pred. No. 8.8e-10;
Matches 66; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1457 GAATTCGAGCTCGGTACCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGC 1516
|||
DB 9 GAATTCGAGCTCGGTACCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGC 68

QY 1517 TGTTTTGGC 1525
|||
DB 69 TGTATAGTC 77

RESULT 14
US-10-013-036-3
; Sequence 3, Application US/10013036
; Patent No. US20020128186A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jeffrey
; TITLE OF INVENTION: Antimicrobial Polypeptide, Nucleic Acid, and Methods of Use
; FILE REFERENCE: 00-1323H
; CURRENT APPLICATION NUMBER: US/10/013,036
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 09/361,900
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: 08/871,924
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Streptococcus mutans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (228)..(782)
; OTHER INFORMATION:
; NAME/KEY: -35 signal
; LOCATION: (177)..(182)
; OTHER INFORMATION:
; NAME/KEY: -10 signal
; LOCATION: (191)..(196)
; OTHER INFORMATION:
; NAME/KEY: RBS
; LOCATION: (218)..(224)
; OTHER INFORMATION:
US-10-013-036-3

Query Match          3.8%; Score 63; DB 12; Length 1323;
Best Local Similarity 100.0%; Pred. No. 2.9e-09;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1454 TAGGAATTCGAGTCGTGACCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTT 1513
|||
DB 912 TAGGAATTCGAGTCGTGACCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTT 971

QY 1514 GGC 1516
|||
DB 972 GGC 974

RESULT 15
US-09-179-536B-267/c
; Sequence 267, Application US/09179536B
; Patent No. US20020042112A1
; GENERAL INFORMATION:
; APPLICANT: Hubert K ster
; David M. Lough
; Guobing Xiang
; TITLE OF INVENTION: DNA DIAGNOSTICS BASED ON MASS SPECTROMETRY
```

```

;
; NUMBER OF SEQUENCES: 320
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/179,536B
; FILING DATE: 26-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/20444
; FILING DATE: 06-NOV-1997
; APPLICATION NUMBER: 08/947,801
; FILING DATE: 08-Oct-97
; APPLICATION NUMBER: 08/933,792
; FILING DATE: 19-Sep-97
; APPLICATION NUMBER: 08/787,639
; FILING DATE: 23-Jan-97
; APPLICATION NUMBER: 08/786,988
; FILING DATE: 23-Jan-97
; APPLICATION NUMBER: 08/746,055
; FILING DATE: 06-No. US20020042112A1-96
; APPLICATION NUMBER: 08/746,036
; FILING DATE: 06-No. US20020042112A1-96
; APPLICATION NUMBER: 08/744,590
; FILING DATE: 06-No. US20020042112A1-96
; APPLICATION NUMBER: 08/744,481
; FILING DATE: 06-No. US20020042112A1-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2004B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 858-450-8400
; TELEFAX: 858-587-5360
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 267:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 103 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: <Unknown>
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 267:
US-09-179-536B-267

Query Match          3.7%; Score 61.6; DB 10; Length 103;
Best Local Similarity 88.2%; Pred. No. 1.7e-09;
Matches 67; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1454 TAGGAATTCGAGTCGTGACCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTT 1513
|||
DB 80 TAGGAATTCGAGTCGTGACCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTT 21

QY 1514 GCGTGTGTTGGCGGAT 1529
|||
DB 20 GGCACGTGGCGTGTGTT 5

Search completed: January 16, 2003, 09:42:30
Job time : 96 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 07:25:01 ; Search time 1687 Seconds
(without alignments)
15869.081 Million cell updates/sec

Title: US-09-830-111A-1
Perfect score: 1653
Sequence: 1 ttttgggggtgaaaagtc.....atgccgaactcagaagtga 1653

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```

EST:*
1:  em_estba:*
2:  em_esthum:*
3:  em_estnu:*
4:  em_estnu:*
5:  em_estov:*
6:  em_estpl:*
7:  em_estro:*
8:  em_esttc:*
9:  gb_estti:*
10: gb_estt2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estom:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	165.4	10.0	561	9 AA901612	AA901612 NCM3A1T3
2	160.4	9.7	656	14 BQ994487	BQ994487 QGF7P20.Y
3	152.8	9.2	843	17 CNS06X9P	AL419459 T3 end of
4	148.4	9.0	690	13 BI722198	BI722198 1031060F0
5	145.2	8.8	986	17 CNS07DTL	AL440911 T7 end of
6	142.6	8.6	621	9 AAS38334	AA538334 VJ03F03.F

7 141.2 8.5 597 14 BQ295774
8 137 8.3 949 9 AL582288
9 136.6 8.3 599 14 BM786570
10 134.2 8.1 897 12 BG252364
11 133.6 8.1 704 13 BJ399049
12 133.2 8.1 735 12 BG125274
13 129.8 7.9 1112 11 AV108392
14 129 7.8 535 12 BF610448
15 125.4 7.6 659 12 BG858032
16 124.8 7.5 485 14 BQ142681
17 123.6 7.5 672 10 AW221898
18 123.2 7.5 650 13 BJ258657
19 120.8 7.3 682 10 AV784831
20 120.6 7.3 570 14 BQ415590
21 120.6 7.3 665 13 BM409824
22 120.2 7.3 769 14 BQ853525
23 120.2 7.3 771 14 BQ853490
24 119 7.2 622 13 BM408807
25 119 7.2 681 14 BM824056
26 117.8 7.1 519 13 BI423817
27 117.2 7.1 455 13 BI708490
28 115.6 7.0 1203 11 AY105886
29 114.4 6.9 601 10 BE432562
30 114 6.9 572 10 BE460050
31 114 6.9 598 14 BQ819659
32 114 6.9 779 14 BQ018783
33 113.6 6.9 574 10 BE436140
34 110.8 6.7 752 14 BQ743828
35 108.4 6.6 558 13 BJ488122
36 108.2 6.5 621 9 AI485604
37 108.2 6.5 629 14 BQ624777
38 107.2 6.5 559 12 BF112700
39 105.8 6.4 984 13 BI756998
40 105.6 6.4 544 10 AW096482
41 105.6 6.4 554 12 BG096598
42 105.6 6.4 658 10 AW093415
43 105.6 6.4 740 14 BQ858515
44 105.2 6.4 676 13 BM139431
45 105 6.4 561 12 BF112719

ALIGNMENTS

RESULT 1
AA901612
LOCUS NCM3A1T3 Mycelial Neurospora crassa
DEFINITION sequence.
ACCESSION AA901612
VERSION AA901612.1 GI:3045589
KEYWORDS EST.
SOURCE Neurospora crassa.
ORGANISM Neurospora crassa
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariales; Sordariaceae; Neurospora.
REFERENCE 1 (bases 1 to 561)
AUTHORS Nelson,M.A., Kang,S., Braun,E.L., Crawford,M.E., Dolan,P.L.,
Leonard,P.M., Mitchell,J., Armijo,A.M., Bean,L., Blueves,E.,
Cushing,T., Erett,A., Fleharty,M., Gorman,M., Judson,K., Miller,R.,
Ortega,J., Pavlova,I., Perea,J., Todisco,S., Trujillo,R.,
Valentine,J., Wells,A., Werner-Washburne,M., Yazzie,S. and Natvig,
D.O.

TITLE Expressed sequences from conidial, mycelial, and sexual stages of
Neurospora crassa
JOURNAL Fungal Genet. Biol. 21, 348-363 (1997)
MEDLINE 97435549
COMMENT Contact: Natvig,D.O./Nelson,M.A.
Department of Biology
University of New Mexico
Catteter Hall, Albuquerque, NM 87131, USA
Tel: 505 277 3411
Fax: 505 277 0304

561 bp mRNA linear EST 10-JAN-2001
AA901612 Mycelial Neurospora crassa cDNA clone NM3A1 5' end, mRNA
BQ295774 1091042F1
AL582288 AL582288
BM786570 K-EST0065
BG252364 602365955
BJ399049 BJ399049
BG125274 EST470920
AV108392 Zea mays
BF610448 NXS_058
BG858032 102405500
BQ142681 123 Metar
AW221898 EST298709
BJ258657 BJ258657
AV784831 AV784831
BQ415590 g24a11.y
BM409824 EST584151
BQ853525 QGB20L24.
BQ853490 QGB20K12.
BM408807 EST583134
BM824056 K-EST0095
BI423817 saf18d11.
BI708490 tt48d07.y
AY105886 Zea mays
BE432562 EST399091
BE460050 EST415342
BQ819659 1030079E0
BQ018783 UI-H-DH1-
BE436140 EST407218
BQ743828 WHE4108.F
BJ488122 BJ488122
AI485604 EST243925
BQ624777 USDA-FP.0
BF112700 EST440290
BI756998 603028715
AW096482 EST289662
BG096598 EST461117
AW093415 EST286595
BQ858515 QGC10H07.
BM139431 Hc_ad 23D
BF112719 EST440309

Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;

Db 340 GAATACGAAAAAATCTTGGTTGGCATTCCAATTAATGATGATGTACTTGATTCACT 399


```
QY 1148 GTCTCGCTACGACCTCGTAAAGCCGCGGTGCAGACCTCCAGCTCGTCTCGCCACC 1207
Db 400 GGCACATCATCTCTAGGGAAGGGTGCATTAATCTGACATTCACCATGAATGTGACA 459
QY 1208 GCCTCGGCGCTCTTCGCATGGAAGCACACGCGGAGCTCGGTCCCATGATCAAGCGCAAG 1267
Db 460 GCTCCAAATATATGCAATGGAGAGTTTCCCGAACTACGTACAGTTGTCGACAGGGC 519
QY 1268 TTCTCTGACCCAGAGACGTGCGAGCGTGCACGGGAGTTGTGCGAGAAAAGTGTGATTG 1327
Db 520 TTGGATGACCCCTGAAAACGTGAATCTTGCTCTGGAGTATCTTGAAAAAGTCNTGGGATA 579
QY 1328 GAGAAGCAGAGACCTTGGCGGAGGAGTATGCCAGAAAGCGTGGATGCAATTCGACG 1387
Db 580 CNAAGAAACAGAGACTGGCAGCAAGCATGCCAGCTGGCATCTGCTGCAATTGATTCA 639
QY 1388 TTCCCGGAGAGT 1399
Db 640 TTGCTGAAAT 651

RESULT 3
CNS06X9P/c
LOCUS
DEFINITION
T3 end of clone AX0AA039F06 of library AX0AA from strain CBS 7064
of Pichia farinosa, genomic survey sequence.
ACCESSION
AL419459
VERSION
AL419459.1 GI:12202637
KEYWORDS
GSS.
SOURCE
Pichia farinosa.
ORGANISM
Pichia farinosa
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Pichia.
1 (bases 1 to 843)
Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Winkler,P. and Weissenbach,J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
FEBS Lett. 487 (1), 3-12 (2000)
20584711
PUBMED
11152876
2 (bases 1 to 843)
de Montigny,J., Spehner,C., Souciet,J., Tekala,F., Dujon,B.,
Wincker,P., Artiguenave,F. and Potier,S.
Genomic exploration of the hemiascomycetous yeasts: 15. Pichia
sorbitophila
FEBS Lett. 487 (1), 87-90 (2000)
20584725
PUBMED
11152890
3 (bases 1 to 843)
Genoscope.
Direct Submission
Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremlieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
Keywords for description of this sequence and for the sequence of
the other extremity of this insert.
FEATURES
Location/Qualifiers
1..843
/organism="Pichia farinosa"
/strain="CBS 7064"
```

```
/db_xref="taxon:4920"
/clone="AX0AA039F06"
/clone_lib="AX0AA"
/note="end : T3"
/note="complement(<309..>839)"
/note="similar to Saccharomyces cerevisiae ORF YBR003w [
COO1 : hexaprenyl pyrophosphate synthetase precursor ]"
/evidence=not_experimental
BASE COUNT 350 a 143 c 139 g 192 t 19 others
ORIGIN
Query Match 9.2%; Score 152.8; DB 17; Length 843;
Best Local Similarity 57.8%; Pred. No. 1.5e-29;
Matches 288; Conservative 2; Mismatches 194; Indels 0; Gaps 0;
QY 965 CAGGAAACGTTTCGATTACTATTTTCAGAAAGACTTACTTGAAGACTCGTCTTGATTGCC 1024
Db 728 CAAGCAGCATTTGAATACTATTTCACAGACTTACTTGAAGAACTCGTCTTTAATGCT 669
QY 1025 AAGTCGTGACAGCAAGTGGCTTCTGGGTGGTCTACGCCGTGAGGTGCTGATGCTGCT 1084
Db 668 AAACCTGTCCAGGCGAGTGCAGTGTTCAGTGGAGCGCAAGATGATGTTATTGGAATTGC 609
QY 1085 TATGCTTACGGAAGGAACCTTGGTTTGGCAATTCAGATCGTCGACGACATGCTCGACTAC 1144
Db 608 TATGAGTTTGGCGGTAAATTTGGGTTTATGCTTCCAAATAGTAGACACATATTAGACTAC 549
QY 1145 ACCGTCTCCGCTACCGACCTCGGTAAGCCGCGGTGCAGACCTCCAGCTCGGTCTCGCC 1204
Db 548 ACCTCAAGCGACAATACTTTTGGTAAACCTAGTCAAGCAGATCTTAAGTTAGTTAGCA 489
QY 1205 ACCGCGCGGCGCTCTTCGCATGGAAGCACCACCGCAGCTCGGTCCCATGATCAAGCGC 1264
Db 488 ACTGCTCCAAATTTTATTCGCTTGAAGAGAGAGCAAACTTGGCCAAATTAATCGCCAGG 429
QY 1265 AAGTTCTCTACCCAGAGACGTCGAGCGTGCACGCGAGTTGCTCGAGAAAGTATGGA 1324
Db 428 AAATTTAGCGAGGAGAGACGTTGAATCGCACTGTAATGCCGTGTAGAAATTTGATGGT 369
QY 1325 TTGGAAGACAGAGAGCCTTGGCGGAGGAGTATGCCAGAAAGCGTTGGATGCAATTCGG 1384
Db 368 TTAGAGAGAGAGAGAAATGGCTGAGGAGTACCGTTTCAAGGCTTTAGAAAACCTGAGA 309
QY 1385 ACCTTCCCGAGAGTCCGGCAGCGAAGGCTTTGAGCAGGTTGAC 1428
Db 308 TTCTTGTGATTCCGWTGCCWTKTATTTCTTTGTTTTTTTAC 265

RESULT 4
BI722198
LOCUS
DEFINITION
1031060E08.y1 C. reinhardtii CC-1690, Stress II (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
BI722198
ACCESSION
VERSION
KEYWORDS
EST.
SOURCE
Chlamydomonas reinhardtii.
ORGANISM
Chlamydomonas reinhardtii.
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
1 (bases 1 to 690)
Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
Unpublished (2001)
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
```

Email: chausereduke.edu.
Location/Qualifiers

FEATURES
source

1. 690
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"

/clone_1lb="C. reinhardtii CC-1690, Stress II (normalized), Lambda Zap II"
/note="Vector: Bluescript II SK-; Site 1: EcoRI; Site 2: XhoI; Stress condition II library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH4+ - containing) and shifted to TAP - NO3- (24hrs); H2 production conditions (0, 12hr, 24hr) see Neils et al., (2000) Plant Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into Lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. Bluescript II SK- plasmids were excised from the Lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

BASE COUNT 113 a 225 c 248 g 104 t

ORIGIN

Query Match 9.0%; Score 148.4; DB 13; Length 690;

Best Local Similarity 55.5%; Pred. No. 2e-28;

Matches 321; Conservative 0; Mismatches 236; Indels 21; Gaps 1;

```

QY 641 GAGGAGACAGCGTGAATATCTGCGCTCGCAACGCGGTTGCTGAGATCAGGAGATG 700
DB 134 GCGGAGACCGCGATGATCGCGCGGCGAGCAGCGCTCGCGAGATCGCGAGACTC 193
QY 701 ATTCATGACAGATACATCTCTCCAGACGACTTATCGACGTTCCGAGACCAAGCAAC 760
DB 194 ATCCACGTGGCGACCTGCTTACGATGATGTGATGATGCGGCGAGACCGCGCGCGC 253
QY 761 GCACCATCCGGAACCAAGCATTCGGAACCAAGATGGCGATTTGGCTGGTATTCTTG 820
DB 254 GTGCTGTGCTCAATACAGGTGGCAACAGACGCGCATTTCTGGCGGGAATTTCTG 313
QY 821 TTGGACGCGGCGTGTGATTCAGTGGCAAGTTGGCAATCCGAGGATGATTGAGCTTTG 880
DB 314 CTGGCGGCGGCTCGTGAAGCTGCGCTGCGCAACAGTGAATCGTGAAGCTGATG 373
QY 881 GCTACTGTATTGCAAACTTGTTGAGGAGATTCATGCACTTGAAATAATCTGTTGAT 940
DB 374 AGTCAGGTGCTGAGACATGTGCTCGGGCGAGATCATGACGATGACCGCCACAGGAG 433
QY 941 GATCGATGAGGCTAGCGGCGACGAGAAAGTTCATTTGAGAAAGACTTAC 1000
DB 434 CAGTCTCTGAC-----TGAAGCACTACCTGGCCAGAACCTTAC 472
QY 1001 TTGAAGCTGCGCTTGTATTGCCAAGTCGTGACAGCAAGTGGCTTTGGGTGCT 1060
DB 473 TGAAGACAGCGACCTCATGCAAAAGCTCGGCTCCGCGGCGTGGCGGAGCG 532
QY 1061 AGCGCTGAGTGTGATGCTGCTTATGCTTACGGAAGAACTTGTGTTGCACTTCCAG 1120
DB 533 GCGCGAGGCTGTGCGCATGCGCGTGAAGCTGACCGCCCACTGGGCGATCGCTTCCAG 592
QY 1121 ATGTCGACGATGCTGATGATACCGCTTCGCTACCGACCTCGGTAAGCCGCCGCT 1180
DB 593 GTGTGATGCTGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 652
QY 1181 GCAGACCTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1218
DB 653 AAGCAGATGCGCAGTGGGCTGGCCACGCGCGGTTGCT 690

```

RESULT 5

CNS07DTL/C

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

CNS07DTL 986 bp DNA linear GSS 08-JUL-2001

T7 end of clone BD0A015F02 of library BD0A from strain CBS 94 of

Candida tropicalis, genomic survey sequence.

AL440911

AL440911.1 GI:12224322

GSS

Candida tropicalis.

Candida tropicalis

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; mitosporic Saccharomycetales; Candida.

1 (bases 1 to 986)

Soulier, J. L., Aigle, M., Artiguenave, F., Blandin, G.,

Bolotin-Fukuhara, M., Bon, E., Broclet, P., Casaregola, S.,

de-Montigny, V., Dujon, B., Durand, P., Lepingle, A., Llorente, B.,

Malpertuy, A., Neuvéglise, C., Ozier-Kalogeropoulos, O., Potier, S.,

Saurin, M., Tekala, F., Toffano-Nicche, C., Wesolowski-Louvel, M.,

Wincker, P., and Weissenbach, J.

Genomic exploration of the hemiascomycetous yeasts: 1. A set of

yeast species for molecular evolution studies

FEMS Lett. 487 (1), 3-12 (2000)

1152876

2 (bases 1 to 986)

Blandin, G., Ozier-Kalogeropoulos, O., Wincker, P., Artiguenave, F. and

Dujon, B.

Genomic exploration of the hemiascomycetous yeasts: 16. Candida

tropicalis

FEMS Lett. 487 (1), 91-94 (2000)

20584726

3 (bases 1 to 986)

Genoscope.

Direct Submission

Submitted (08-SEP-2000) Genoscope - Centre National de Sequençage,

2 rue Gaston Crémieux, CP 5706, 91057 Evry cedex, FRANCE. (E-mail :

seque@genoscope.cns.fr - Web :

This GSS is part of a random genomic sequencing program of thirteen

yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces

exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,

Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces

lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia

angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,

Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to

5 kb were prepared and both extremities were sequenced. See

keywords for description of this sequence and for the sequence of

the other extremity of this insert.

Location/Qualifiers

1. 986

/organism="Candida tropicalis"

/strain="CBS 94"

/db_xref="taxon:5482"

/clone="BD0A015F02"

/clone_1lb="BD0AA"

/note="end : 77"

misc_feature

complement(1<6..>899)

/note="similar to Saccharomyces cerevisiae ORF YBR003w [

COQ1 : hexaprenyl pyrophosphate synthetase precursor]"

evidence-not experimental

BASE COUNT 298 a 209 c 141 g 337 t 1 others

ORIGIN

Query Match 8.8%; Score 145.2; DB 17; Length 986;

Best Local Similarity 52.2%; Pred. No. 1.9e-27;

Matches 458; Conservative 0; Mismatches 318; Indels 102; Gaps 2;

```

QY 671 CAACGCGGTGGCTGATGATCAAGAGATGATGATGATGATGATGATGATGATGAT 730
DB 887 CAAAGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 828
QY 731 GTTATGAGCGTTCCGAGACGAGGAAACGACATCCGGAACACGAGCATTCGGAAC 790
DB 827 GTGATGATTTGCTGATTCAGAAAGGTAGACCAAGCGGAAATATTCATTTACAAAT 768

```



```

Db      475 GAGATGTGGACAGACGACGATGATATTACAGAGTGTGGCTGACGAAACAACCT 534
Qy      1341 CTTTGGCGGAGAGTATGCCGAGAGCGTTGATGCAATTCGACCTCCCGAGAGTC 1400
Db      535 ACCTCGCCAGACATCTACGCCACAAGCTGTGAGAGATCAGAGACCTTAGACATCTA 594
Qy      1401 CGGCACGGAAGCTTTGGAGCAGTT 1425
Db      595 CAGAAAGGAGCGCCCTCATTCAGCT 619

RESULT 7
LOCUS   B0295774
DEFINITION 1091042F12.v1 1091 - Immature ear with common ESTs screened by
          B0295774
          Schmidt lab Zea mays cDNA, mRNA sequence.
ACCESSION B0295774
VERSION   B0295774.1
KEYWORDS  GI:20811296
SOURCE    EST.
          Zea mays.
          Zea mays.
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
          clade; Panicoidae; Andropogoneae; Zea.
          1 (bases 1 to 597)
REFERENCE 1
AUTHORS   Walbot,V.
TITLE      Maize ESTs from various cDNA libraries sequenced at Stanford
          University
JOURNAL   Unpublished (1999)
COMMENT   Contact: Walbot V
          Department of Biological Sciences
          Stanford University
          855 California Ave, Palo Alto, CA 94304, USA
          Tel: 650 723 2227
          Fax: 650 725 8221
          Email: walbot@stanford.edu
          Plate: 1091042 row: F column: 12.
          Location/Qualifiers
            1..597
              /organism="Zea mays"
              /cultivar="OH43"
              /db_xref="taxon:4577"
              /clone_lib="1091 - Immature ear with common ESTs screened
              by Schmidt lab"
              /cissue_type="Inflorescence meristem - floral organ
              primordia"
              /dev_stage="0.5 cm to 2 cm"
              /lab_host="Stratagene XLDR"
              /note="Organ: Immature ear; Vector: PAD-GAL4; Site_1:
              EcoRI; Site_2: XhoI; RNA from library 606 was filtered for
              common ESTs found in 606."
BASE COUNT 155 a 125 c 148 g 169 t

Query Match 8.5%; Score 141.2; DB 14; Length 597;
Best Local Similarity 55.9%; Pred. No. 1.5e-26;
Matches 322; Conservative 0; Mismatches 233; Indels 21; Gaps 2;

Qy      670 GCAACGGCGGTTGGCTGATCAGGAGATATCCATGACATCACTCCCTCCAGACGA 729
Db      3 GCAACGAAACATTCGTGATGATCACTGAATGATTCAGTCCGAAAGCTTCGTGATGATGA 62
Qy      730 CGTATGCAACCTTCGAGACGAAAGCAACATCCGAAACGAGCATTCGGA 789
Db      63 TGTCTTGATATGCTGATCTAGGCGGTGCTGATTCATGATCTCATCAGGGGA 122
Qy      790 CAAGATGGCAATTTTGGCTGTGATTTCTTTGGGAGCGGCGTGTGTGATTGGCGAG 849
Db      123 CAACCTTTCTGTGCTGGCTGTGATTTCTTCTGTGAGACATGTGTGGCCCTTGGAGC 182
Qy      850 GTTGGCAATCCGAGGATGATGAGCTTTGGCTAAGCTTTTGGCAACTTGTGGAGG 909

```

```

Db      183 ACTGGGAACACAGAGGTGTTCTCTAATGGCACTGCTGAGAACATCTAGTACTGG 242
Qy      910 AGAGTTTCATGCACTGTAATAATCTGTTGATGATGCGATTGAGCTACGGCAGCAGGA 969
Db      243 TGAACCTATGAGAGATCTCAACAAGCAGAGA-----GCAGCGGAGGAGCATG- 288
Qy      970 AACGTTCGATTCCTAATTTGCAAGAAGCTTACTTGAAGACTGCTGATTTGCCAATGC 1029
Db      289 -----GAGTACTACTGACAGAGACATCTACAAAACGGCATATTGATATCAAAATG 341
Qy      1030 GTGACAGCAAGTGGCTTCTGGTGTGCTACGCTGAGGTGCTGATGCTTATGCTATGAC 1089
Db      342 TTGCAAGGCTGTGCTATTTCTTGACAGGACACAATGAGTCTGGTCTTGCAATATGA 401
Qy      1090 TTACGAGGAACCTTGTGTTGGCATTCGATCCTGACAGCATGCTGCACTACACCTG 1149
Db      402 ATATGCTGCAAAACCTGGGCTGACCTCCAGATTAATGATGATGTTGATTTGACCGG 461
Qy      1150 CTCGGCTACGACCTCGGTAGACCGCGCGGTGACAGCTCCAGCTCGGTCTGCGACCGC 1209
Db      462 CACCTTGATCCCTTGGGAGGGTTTCATTGTGATATTCGCCACGAAATTAATTAATTCG 521
Qy      1210 GCCGCGCTCTTGCATGAGACACACCGCGAGCT 1245
Db      522 CCGAGCTATATGCGATGAGGAATTCACCACT 557

RESULT 8
LOCUS   AL582288/c
DEFINITION AL582288 LTI_NFL010_BC2 Homo sapiens cDNA clone CS0D005YM10 3
          prime, mRNA sequence.
ACCESSION AL582288
VERSION   AL582288.1
KEYWORDS  GI:12950123
SOURCE    EST.
          human.
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
          1 (bases 1 to 949)
REFERENCE 1
AUTHORS   Li,W.B., Gruber,C., Jesssee,J. and Polayes,D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
          Genoscope - Centre National de Sequencage
          BP 191 91006 EVRY cedex - France
          Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
          Location/Qualifiers
            1..949
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="CS0D005YM10"
              /clone_lib="LTI_NFL010_BC2"
              /sex="male"
              /cissue_type="B cells from Burkitt lymphoma"
              /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
              was primed with a NotI-OLIGO(dT) primer. Five prime end
              enriched, double-stranded cDNA was digested with Not I and
              cloned into the Not I and Eco RV sites of the pCMVSPORT 6
              vector. Library was normalized. Library was constructed by
              Life Technologies. Contact : Feng Liang Life Technologies,
              a division of Invitrogen 9800 Medical Center Drive
              Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
              Email : filiang@life.com URL :
              http://fulllength.invitrogen.com"
BASE COUNT 285 a 192 c 185 g 283 t 4 others

Query Match 8.3%; Score 137; DB 9; Length 949;
Best Local Similarity 51.5%; Pred. No. 2.8e-25;
Matches 354; Conservative 1; Mismatches 311; Indels 21; Gaps 1;

Qy      781 ATTGGAAACAGAGTGGCAATTTGGCTGTGATTTCTTTGGGAGCGGCGTCTGTTCG 840

```

Db	945	ATCTGGTGAAGAGCTGTTCTTCTGGAGATTAAATTTCTTCTCCAGCATCTATAGC	886
Qy	841	ATTGGGAGGTGGCGCAATCCGAGGATGATGAGCTTTTGGCTACTGTTATTGCAAACTT	900
Db	885	TCGGCAGCAATTTGGAATAACAACCTGTTATCTATTATTAACCAAGTTATTGAAGATTT	826
Qy	901	GGTTGAGGAGAGTTCATGTCAGTGTGAATAACTGTTGATGTCGATTCGAGGCTACGGC	960
Db	825	GGTGGTGGTGAATTTCTTCAGCTCGGGTCAAAAGAAATGA	784
Qy	961	GACGCAAGAAACGTTGCTATTACTATTTCGAGAAAGCTTACTTGAAGACTCGCTCTTGAT	1020
Db	783	---GAATGAAGATTGACACATCTTTCAGAGACATTCAAGAGACCCGACCTGAT	727
Qy	1021	TGCCAGTCTGTCAGAGCAAGTGGCTTCTGGTGGTGGTTCAGCGCTGAGTTCGATGC	1080
Db	726	AGCCAACAGTTGTAAGACAGTCTCTGTTCTAGGATGTCGCGACCCAGTGGTGCATGAT	667
Qy	1081	TGCTTATGCTTACGAGGAGCACTTGGTTGGCATTCAGATGTCGACGACATGCTCGA	1140
Db	666	CGCTATCAGTACGGAATAATGTAGGAATAGCTTTTCAGCTAATAGATGATGATTGGA	607
Qy	1141	CTACACGCTCCGCTACCGACCTCGGTAAGCCGCGGTGCAGACCTCCAGCTCGGTCT	1200
Db	606	CTTCACTCTGTTCTGACAGATGGCAACCAACATCAGTGTGATCTGAAGCTCGGTT	547
Qy	1201	CGCACGCGCGCGCCCTCTTCGATGGAAGACCAACCGCGAGCTCGGTCCCATGATCAA	1260
Db	546	AGCCACTGGTCTGCTCTGTTTGGCTGTAGCAGTTCGCCAGAAATGAATGCTATGATCAT	487
Qy	1261	GCAGAGTTCTTGACCCAGAGAGCTGAGCGTGCACGGAGTTGGTCGAGAAAGTGA	1320
Db	486	GCACCGGTTTCAGTTTGGCTGAGATGTAGACAGAGTTCGACGATGATCTACTACAGATGA	427
Qy	1321	TGATTGGAGAAGACGAGAGCCCTTGGGGAGGAGTATGCCAGAGGCGTTGGATGCAAT	1380
Db	426	TGTTGTGCAACAACACCTTCCTCGCCAGAGAGCTGCCATGAAGCAATAGAGAGAT	367
Qy	1381	TCGGAGCTTCCCGGAGAGTCCGGCAGGAAGGCTTTGGAGCAGTTGACGGACAGGTGTT	1440
Db	366	CAGTAAACTTCGACCATCCCGCAAGAAAGAGATGCCCTCATTCAGCTTTCAGAAATGTACT	307
Qy	1441	GACTAGTCAAGATAGAAATTCAGCT	1467
Db	306	CACAAGAGATAAATGACAACTTTCT	280

RESULT 9

BM786570

LOCUS

DEFINITION

BM786570

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

599 bp mRNA linear EST 05-MAR-2002

K-EST0065318 S9SNU601 Homo sapiens cDNA clone S9SNU601-39-C05 5', mRNA sequence.

BM786570

BM786570.1 GI:19134802

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 599)

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 39	row: C	column: 05
High quality sequence stop: 599.	Location/Qualifiers	
1. 599	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="S9SNU601-39-C05"	
	/clone_lib="S9SNU601"	
	/sex="M"	
	/tissue_type="Ascites"	
	/cell_type="Epithelial"	
	/cell_line="SNU-601"	
	/lab_host="Top10F"	
	/note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI; Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then dephosphorylated with tabacco acid pyrophosphatase (TAP). The dephosphorylated intact mRNA was ligated with DNA-RNA linker including SfiI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized with Superscript II using SfiI oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with SfiI and cloned into DraIII- digested pME18-FL3 vector. The obtained cDNA vectors were used for transformation of competent cells E. coli top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."	
BASE COUNT	176 a 124 c 142 g 157 t	
ORIGIN		
Query Match	8.3%; Score 136.6; DB 14; Length 599;	
Best Local Similarity	53.2%; Pred. No. 2.6e-25;	
Matches 329; Conservative	0; Mismatches 269; Indels 21; Gaps 1;	
Qy	738	ACGCTTCCGAGACACGACGACACCATCCGAAACGAGCATTCGGAACAGATGG 797
Db	1	ACGATCAAGTTCTCGAAGAGGAAAAACACACAGTTAATAAGATCTGGGTGAAAGAGG 60
Qy	798	CGATTTTGGCTGCTGATTTCTTCTGGGACGGCGCTCTGTTGCATTGGCGAGGTGCGCA 857
Db	61	CTGTTCTTCTGGAGATTAAATCTTTCACAGCATCTATAGCTTCGGACAGATTGGAA 120
Qy	858	ATCCGAGGTGATGAGCTTTTGGCTACTGTTATTCGAACTTGGTTGAGGAGAGTTCA 917
Db	121	ATACAACTGTTATATCTATTTTAAACCAAGTTATTGAAGATTTGGTGGGTGAAATTC 180
Qy	918	TGCAGTTGAAAATACTGTTGATGTCGATTGAGCTACGGCAGCAGCAAGACGTTCCG 977
Db	181	TTCAGCTCGGGTCAAAAGAAAATGA-----GAATGAAGATTG 219
Qy	978	ATTACTATTTCGAGAAGACTTACTTGAAGACTCGCTCTTGTGTTGCAAGTTCGTGCAGAG 1037
Db	220	CACACTACCTTNGAAGAGATTCAGAAAGACCCGACGCTGATAGCCACAGTTGTAAG 279
Qy	1038	CAAGTCGGCTTCTGGTGGTGTACCCCTGAGGTGCTGATGCTGCTGTTATGTTACGGAA 1097
Db	280	CAGTCTCTGTTCTAGGATGTCCCGACCCAGTGGTGCATGAGATCGCTATCAGTACGGAA 339
Qy	1098	GGAACTTGGTTGGATTCAGATCGTCGACGACATGCTCGACTACACGCTCTCCGCTA 1157
Db	340	AAAATGTAGGAATAGCTTTTCAGCTAATAGATGATGATTATGGACTTTCACCTCGTGTCTG 399
Qy	1158	CGACCTCGTGAAGCCCGCGGTGCGAGACTCCAGCTCGGTCTCGCCACCCGCGCGGCC 1217
Db	400	ACCAGATGGGCAACCAACATCAGCTGATCTGAAGCTCGGGTTAGCCACTGGTCTGTGCC 459
Qy	1218	TCTTCGATGGAGACACCGCGAGCTCGGTCCCATGATCAAGCGCAAGTCTCTGACC 1277
Db	460	TGTTTGCCTGTGACGAGATTTCCAGAAATGAATGCTATGATCATGCGACGCTTCAGTTGC 519
Qy	1278	CAGGACAGCTCGAGCGTGCACCGAGTTGGTTCGAGAAAGTGTATGATTTGGAGAACGA 1337

```

Db 520 CTGAGATGTAAACAGAGCTCGACAGTATGTACTACAGAGTGTGTGTGCAACAAACA 579
Qy 1338 GAGCCTTGGCGGAGAGTA 1356
Db 580 CCTACCTCGCCGACGAGTA 598

RESULT 10
Bg252364 897 bp mRNA linear EST 13-FEB-2001
LOCUS 602365955F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4474188 5'
DEFINITION mRNA sequence.
ACCESSION BG252364
VERSION BG252364.1 GI:12762180
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 897)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapds-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM10297 row: n column: 13
High quality sequence stop: 754.
location/Qualifiers
1. 897
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4474188"
/clone_1ib="NIH MGC 90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 255 a 194 c 197 g 251 t
ORIGIN
Query Match 8.1%; Score 134.2; DB 12; Length 897;
Best Local Similarity 51.5%; Pred. No. 1.5e-24;
Matches 350; Conservative 0; Mismatches 308; Indels 21; Gaps 1;

Qy 768 CCGGAAACGAGCATTGCGAAACAGATGCGCATTTGGCTGTGATTTCTTGGGAC 827
Db 17 CAGTTAATAGATCTGGGCTGAAAGAGGCTGTTCTTGGAGATTAATTTCTTG 76
Qy 828 GGGCGCTTGTGATTGGGAGGTTGGCAATCGGAGGATGAGCTTTGGCTACG 887
Db 77 CAGCATCTAAGCTCTGGCAGCATGGAATGAAATCACTGTATATCTAATTTAACCAAG 136
Qy 888 TTATTGCAACTTGTGTGAGGAGAGTTCATGCAATTGAAAAATACGTTGATGACGA 947
Db 137 TTATTCAGATTTGGTGGCTGGGATTTCTTACGTCGGGTAAAGAAAAATGA----- 191
Qy 948 TTGAGGCTACGGGACGAGAAACGTTGATTACTATTGTCAGAAAGCTTACTTGAAG 1007
Db 192 -----GAATGAAAGATTGGACACACTACCTTGAGAGACATTCAGAAAGA 235
Qy 1008 CTGCGCTTGTGATGCGAAGTCTGACAGCAAGTGGCTTCTGGGTGGTGTGACGCTG 1067
Db 236 CCGCAGCCTGATGACCAAGTTGTAAGAGCTCTCTGTCTTAGAGATGCCGACCAG 295

```

```

Qy 1068 AGTTGCTGATGCTGCTTATGCTTACGGAAGAACTTGTTGGCATTCAGATCGTC 1127
Db 236 TGGTCATGAGATCGCCCTATAGTACGAAAAAATGTAAGAAATGCTTTTACGTAATAG 355
Qy 1128 ACAGATGCTCGACTACACACCGCTCCGCTACCGACCTCGGTAAAGCCGCGGTGACAC 1187
Db 356 ATGATGTATTGGACTTTCACCTCGTGTGACCATGATGGCAAAACATCAGCTGATC 415
Qy 1188 TCCAGCTCGTCTCGCCACCGCGCGGCTTTCGATGGAAGACACCGCCGAGCTCG 1247
Db 416 TGAAGCTCGGGTTACGACCTGCTGCTGCTGTTGCTGTCAGAGATTCGCAAAATGA 475
Qy 1248 GTCCATGATCAAGGCAAGTTCTGTACCCAGAGAGAGTGTGAGCTGACCGAGTTGG 1307
Db 476 ATGCTATATATCATGGAAGGTTGCTGAGATGTAGACAGAGCTGACAGATAG 535
Qy 1308 TCGAAGAAATGATGATTTGGAGAAACGAGAGGCTTGGCGGAGAGATATGCCAGAG 1367
Db 536 TACTACAGATGATGTGTGCAACAAACATTACTCCCGACAGTACTGCTATGAAGC 595
Qy 1368 CTTGATGCAATTGGAAGTTCGCCGAGAGTCCGCGACGGAAGGCTTTGAGCAGTTGA 1427
Db 596 AATAAGAGAGATCACTAATTAAGTTCAGCATCCCGCAAGAAAGATGCCCTCATTAAGCTT 655
Qy 1428 CCGCAGAGGTGTACTTG 1446
Db 656 TCAGAAATGTACTCACAAG 674

RESULT 11
Bg399049 704 bp mRNA linear EST 10-MAR-2002
LOCUS Bg399049 Dictyostelium discoideum cDNA library, SF Dictyostelium
DEFINITION dictyostelium cDNA clone dds3h23 3', mRNA sequence.
ACCESSION Bg399049
VERSION Bg399049.1 GI:19311966
KEYWORDS EST.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum.
Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
1 (bases 1 to 704)
REFERENCE Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the slug stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yatae, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tehini@genes.nig.ac.jp.
location/Qualifiers
1. 704
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dds3h23"
/clone_1ib="Dictyostelium discoideum cDNA library, SF"
/sex="mat A"
/dev_stage="slug stage"

BASE COUNT 233 a 125 c 91 g 255 t
ORIGIN
Query Match 8.1%; Score 133.6; DB 13; Length 704;
Best Local Similarity 51.0%; Pred. No. 1.8e-24;
Matches 365; Conservative 0; Mismatches 344; Indels 27; Gaps 1;

Qy 749 ACCACAGAAAGCAGATCCGGAACACGAGCATTCGAAACAGATGCGATTTGGCT 808
Db 699 ACTAGAGAGATGATATGATTAATCATATGTTTACAAATTAATTTGGCAATTTTATGT 634
Qy 809 GGTATTTCTTGTGGAGCGGCGTCTGTGATTTGGCAGAGTTGCCAATCCGAGGTG 868

```

```
Db 633 GGTGATTATTATTTGGCAAGCAACAGTAGTTTATCAAAATAGAAATCCAGATGTA 574
QY 869 ATTGAGCTTTTGGCTACTGTTATTGCAAACTTGTTGAGGAGAGTTTCATCGAGTTGAAA 928
Db 573 ACTGAATGATGATCAACTGCTGGCCGAGTTGTTGAGGTTGAATTTATGCAAGCAAAA 514
QY 929 AATACGTTTCATGATGATGAGGCTACGGCAGCGAGAAAGCTTCGATTACTATTG 988
Db 513 TCAAAATGGTCTGCTCTTTT-----GATAATTATTA 481
QY 989 CAGAAGACTTACTTGAAGACTCGCTGCTGATTGCCAAGTCGTGAGCAAGTGGCCTT 1048
Db 480 CAAAAACCTTATTAAACCTGGTCTTTTAAATTAACAACAGTTGTAGATCTGCTGCTATC 421
QY 1049 CTGGTGGTCTAGCGCTGAGGTTGCTGCTGCTTATGCTTACGGAAGAAACCTTGGT 1108
Db 420 CTTTCTGGTCAGATTCAAAACATTAATAATTTCAACTGAATTTTGGAAAGAAATTTAGGT 361
QY 1109 TTGGCATTCCAGATCGTCGAGCATCTCGACTACACGCTCTCCGCTACGACCTCGGT 1168
Db 360 TTAGCTTTTCAAAATGTTGATGATTTTATTAGATTATACAGGATCAGCTGAAGAATGTGGA 301
QY 1169 AAGCCCGCGGTGCAGACCTCCAGCTCGGTCTGCGCACCGCGCGCCCTCTTCGCAATGG 1228
Db 300 AAGCTACATCAGTTGATTTAACTTTGGGATTAGCAACAGACACAGTTTATATGCAACC 241
QY 1229 AAGCACACCCGAGCTCGGTCCCATGATCAAGCGCAAGTTCTCTGACCCAGGAGAGCTC 1288
Db 240 CAAGAATTCCACAAATFAGAGAAATTAATTAAGAAAGATTTTCAGAGATTGGTGATTT 181
QY 1289 GAGGTGCAGCGAGTTGGTCGAGAAAGTGTGATGAGTGGAGAGAGAGAGCTTGGCG 1348
Db 180 GAAGAAGCAAGAGGTTGGTTGCTTTAAGTAAAGGTATCGAGAAATCTAGAAATTTAGCT 121
QY 1349 GAGGAGTATCCAGAGAGGCTTGGATGCAATTCGGAGCTTCCGGAGAGTCCGGCAGCG 1408
Db 120 ATGAATATTGTAATCTGCAATTCATCTTATTAATAATTAACCAATCTGAATCAAGA 61
QY 1409 AAGGCTTTGGAGCAGTTGACGGCAAGGTTGCTAGTGGTCAAGATAGGAATTCGA 1464
Db 60 GATCTTTTATCATATCTCATATGTTGTTTACAAGACAAAGTAAATACGA 5

RESULT 12
BG125274
LOCUS EST470920 tomato shoot/meristem Lycopersicon esculentum cDNA clone
DEFINITION CTOF8K7 5' sequence, mRNA sequence.
ACCESSION BG125274
VERSION BG125274.1 GI:12625462
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 735)
AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T.,
Hansen,C., Roming,C. and Tanksley,S.
TITLE Generation of ESTs from tomato shoot/meristem tissue
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
1. .735
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"

FEATURES
source
```

```
/clone="CTOF8K7"
/clone_lib="tomato shoot/meristem"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."
BASE COUNT 200 a 149 c 168 g 218 t
ORIGIN
Query Match 8.1%; Score 133.2; DB 12; Length 735;
Best Local Similarity 59.1%; Pred. No. 2.4e-24;
Matches 259; Conservative 0; Mismatches 158; Indels 21; Gaps 1;
QY 682 GGCTGAGATCAGCGAGATGATCCATGCAGCATCACCTCTCCAGACAGACGCTTATCGACGC 741
Db 278 GGCTGAGATCAGCTGAGATGATCCATGTTGCTAGCCTACTTCATGATGATGATGATGATGA 337
QY 742 TTCGAGACAGCAAGAACCCATCCGGAAACACGAGCATTCGGAAACAAAGATGGCGAT 801
Db 338 TGCTGACACAAAGACGTGGGATAGTCTTTTAAACTTTTGATGGAAATTAAGCTAGCTGT 397
QY 802 TTTGGCTGCTGATTTCTTGTGGACGGGCTCTGTTCGATTGGCGAGGTTGGCAATCC 861
Db 398 ACTAGCCGGAGACTTTTGTCTTTCCCGAGCATGTGTGGCACCTTGCCTCTTGAAGAACAC 457
QY 862 GGAGGTGATTGAGCTTTTGGCTACTGTTATTGCAAACTTGGTTGAGGAGAGTTTCATGCA 921
Db 458 AGAGGTGTTATGTTCTTCTGGCAACTGTTGTGGAAACATCTTGTACTGGAGAGACAATGCA 517
QY 922 GTTGAATAAATACTGTTGATGATCGGATTGAGGCTACGGCAGCAGCAGGAAACGTTTCGATTA 981
Db 518 AATGACGACTTCTTCTGATGA-----ACGTTGTAGCATGGAGTA 556
QY 982 CTAATTTGCAAGACACTTACTTTGAAGACTGGCTCTTGTATTGCCAAGTCGTGCGAGCAAG 1041
Db 557 TTATATGCAAGAAACATATTTACAAGACTGTCATCATTTGATTTCAATAGGTGCAAGCAAT 616
QY 1042 TGGCTTCTGGGTGGTGTCTACGCTCAGGTTGCTGATGCTGCTTATGCTTACCGAAGGAA 1101
Db 617 TGCACTACTTGTGGGCTAGTGTCTGAAAGTCTCCGCTGGCTTTTGTACTACGGGAAAAA 676
QY 1102 CTTGGTTTGGCATTCGA 1119
Db 677 TCTGGGATTGGCATTTCA 694
RESULT 13
AY108392
LOCUS AY108392 1112 bp mRNA linear HTC 25-MAY-2002
DEFINITION Zea mays PCO073322 mRNA sequence.
ACCESSION AY108392
VERSION AY108392.1 GI:21211470
KEYWORDS HTC.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1112)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
JOURNAL Overgo Probes
REFERENCE Unpublished (2002)
AUTHORS Coe,E.C.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
Location/Qualifiers
```


source 1. 1112
 /organism="Zea mays"
 /db_xref="MaizeDB:634028"
 /db_xref="taxon:4577"
 /clone="PC007332"
 /clone_1ib="Maize Mapping Project/Dupont Cornsensus library"
 /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

BASE COUNT 344 a 221 c 248 g 299 t
 ORIGIN

Query Match 7.9%; Score 129.8; DB 11; Length 1112;
 Best Local Similarity 49.7%; Pred. No. 2.6e-23;
 Matches 378; Conservative 0; Mismatches 362; Indels 21; Gaps 1;

QY 671 CAACGCGGTGGCTGATGATCAGGAGATGATCAGATCAGATCCTCCACGACGAC 730
 DB 18 CATCGACGGTGGCAGAGATTAATCGAGATGATTCACACTGCGATTAAATACATGATGAT 77
 QY 731 GTTATGAGCTTCCGAGACCAAGCAACGCCATCCGGAACAGGATTCGGAAC 790
 DB 78 GTCATGATGATAGTGGAGAGAGAGAGAGAGAACTATTCACCAACTATATGATGAT 137
 QY 791 AAGATGCGATTTGGCTGGTGAATTTCTTTGGAGAGCGGCGCTGTGGATTCGCGAG 850
 DB 138 CGGTGGCTGATCTGCTGATTTTATGTTTGACCAATCTTCTGGTTCTTGCAAC 197
 QY 851 TTGCGCAATCCGAGAGTATGAGCTTTGGCTACTGTTATTCGAACTTGGTTGAGGA 910
 DB 198 CTAGAAATATTAAGATTAAATGATCAGTCAGTCAATCAAGAGCTTTGCAACGCGC 257
 QY 911 GAGTTCATCAGTGAATAATAGTGTGATGAGGATTCGAGGCTAAGGAGAGAGGA 970
 DB 258 GAGATTAACAGCTTCCACTCTTTTGAATGACATC----- 296
 QY 971 ACCTGATTAATTTTGGAGAGACTTCTGAGAGCTGCTGCTTGGATTCGCAAGTCG 1030
 DB 297 ACATGAGAGACTACTCTTCTCAGAGAGTACTACAGACTGATCTTTGATTCAGGACG 356
 QY 1031 TCGAGAGCAAGTGGCTTCTGAGTGTGCTACGCTGAGAGTTCGATGCTGTTATGCT 1090
 DB 357 ACAAAATCAGCTTCCATATTCAGTGGCTGACAGCACCACTATTTGAAAAAATGTTATGA 416
 QY 1091 TACGGAAGAACTTGGTTTGGATTCAGATGCTGAGAGACATGCTCCAGTACACCGTC 1150
 DB 417 TATGGAGGAATTTGGTCTATCTTCCAGAGTGTGATGACATCTGATTTTACCCAG 476
 QY 1151 TCGGTAACGACCTCGGTAAAGCCGCGGTGACAGACTTCCAGCTCGGTCTGCGCACCGG 1210
 DB 477 TACGCGAAACAATTGGCAAAACAGAGCAAGTGAATTCGCAAGGAAACCTGATGCT 536
 QY 1211 CCGGCGCTTCTGCAATGGAAGACCAAGCCGAGCTCGTCCATGATCAAGCCGAATTC 1270
 DB 537 CCAAGTACTTCTGCTTGAAGTGAACAGAGCTTAAGGAGATCAATGATTTCTGAGTTC 596
 QY 1271 TCTGAGCCAGAGAGAGTCCAGAGCTGAGAGTGTGAGAAAGTATGATGATGAG 1330
 DB 597 AGTGAACAGGATTCGTTAGTACTGCAATAGAGCTCGTTCATGAAGTGGTGGATACG 656
 QY 1331 AAGAGAGAGCTTGGCGGAGAGTATGCCAGAGGCGTTGATGCAATTCGAGAGTTC 1390
 DB 657 AAGGCAATGAGTTCGAGAGAGAGAGAGGAGTTCGCAATTCGAGAGTTCGAGTTC 716
 QY 1391 CCGGAGAGTCCGAGAGAGAGGCTTTGAGACAGTTCAGGA 1431
 DB 717 CCAAGAAAGTACTTCAGAAAGTCCCTTGAAGAAATGATGATGGA 757

RESULT 14
 BE610448
 LOCUS 535 bp mRNA linear EST 14-DEC-2000
 DEFINITION NXSI_058_F04_F NXSI (Nsf Xylem Side wood inclined) Pinus taeda cDNA
 clone NXSI_058_F04_5', mRNA sequence.
 ACCESSION BE610448
 VERSION BE610448
 KEYWORDS EST.
 SOURCE 1010101 pine.
 ORGANISM Pinus taeda
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
 Sederoff, R.
 REFERENCES 1 (bases 1 to 535)
 TITLE Molecular Basis of Wood Formation in the Pine Megagenome
 JOURNAL Unpublished (2000)
 COMMENT Contact: Johnson, Arthur
 North Carolina State University
 Tel: 919 515 7801
 Fax: 919 515 7801
 Email: ajohnson@unity.ncsu.edu
 Seq primer: 73.
 Location/Qualifiers

FEATURES

source 1. 535
 /organism="Pinus taeda"
 /strain="Coastal plain loblolly pine from North Carolina"
 /db_xref="taxon:3352"
 /clone="NXSI_058_F04"
 /clone_1ib="NXSI (Nsf Xylem Side wood inclined)"
 /tissue_type="Xylem"
 /cell_type="side"
 /dev_stage="juvenile"
 /lab_host="XLI-Biue"
 /note="Vector: Bluescript SK; Site 1: Eco RI; Site 2: XhoI
 ; The library is from early (spring) wood, taken from
 three six-year old trees (three different genotypes), in
 the juvenile phase. These trees were induced to form side
 wood by bending to a 45 degree angle and tying them to the
 ground. Differentiating xylem was harvested from the sides
 of the inclined stems, and a mixture of all three
 genotypes was used for the library. oligo-dT primed cDNA
 was directionally cloned into the EcoRI-XhoI Bluescript SK
 vector arms. NOTE: The sequences contain a 'cDNA adapter'
 between the EcoRI site and the start of the EST. The
 adapter sequence is 'AATTCGACAG'."

BASE COUNT

148 a 96 c 131 g 160 t

Query Match 7.8%; Score 129; DB 12; Length 535;
 Best Local Similarity 54.3%; Pred. No. 2.6e-23;
 Matches 298; Conservative 0; Mismatches 230; Indels 21; Gaps 1;

QY 744 CCGAGACAGAGAAACCAACCATCCGGAACAGAGCTTGGAAACAAGATGGGATTT 803
 DB 7 CAGATACCCGTCGAGGTGTGGCTGTTGAATTTGTAATAGGAATTAAGCTTGCTGAC 66
 QY 804 TGGCTGATTTCTTGTGGAGCGGCGTGTGCAATGCGAGGTTGGCAATCCG 863
 DB 67 TGGCTGATTTCTTGTGGAGCGGCGTGTGCAATGCGAGGTTGGCAATGAG 126
 QY 864 AGGTATGAGTTTGGCTACTGTTATTCGAAACTGTGAGGAGAGTTCATGAGT 923
 DB 127 AGTGTGAATTAATTTGTCGAGGTTGGAACATCTTGTCATGTAAGTAATGCAAA 186
 QY 924 TGAATAATCTGTTGATGATGATGATGAGCTACGAGCAGAGAAAGCTTCATTA 983
 DB 187 TGACAAATTAACCTGAAC-----GTGCTGATGATGATGATTT 225
 QY 984 ATTTCAGAGAGCTTACTTGAAGAGTGGTCTTGTGATTCGCAAGTCTGAGAGCAAGT 1043
 DB 226 AATTCAGAGAGCTTCTTGAACAAACAGCATCATGATGCTTAACAGTTCGAAAGGCAATG 285
 QY 1044 CGCTTCTGGGTGTCTACGCGCTGAGGTTGCTGATGCTTATGCTTACGGAAGGAACC 1103

Db 286 CTCTTATGCGGTGACCGAGAGAGTTTGCAATGCTTGCTATGACTATGGAAGAATT 345
 QY 1104 TTGGTTTGGCAATTCAGATCGTCCGACACATGCTCGACTACACCGTCTCCGCTACCGACC 1163
 Db 346 TGGGATTAGCATATCAGTCTAGTTGATGATGCTTACTGCTTACTGCTGACCAACAGCTTCAC 405
 QY 1164 TCGGTAAGCCCGGTGACAGCTCCAGTCTCGTCTCGCCACCGCGCGCCCTCTTCG 1223
 Db 406 TTGGCAAGGGCCCTCTTCTGACATACGCCAGGGGATTTGTTACTGCTCCAATATTGTTG 465
 QY 1224 CATGGAAGACACCGCGAGCTCGGTCCCATGATCAAGCGCAAGTTCTCTGACCCAGGAG 1283
 Db 466 CTTTGAAGAGTTTCCACAGTTGCATGATGATATCAGCCGAAAGTTCAAAAAGCCTGGAG 525
 QY 1284 ACCTCGAGC 1292
 Db 526 ATATTGATC 534

RESULT 15
 BG858032
 LOCUS
 DEFINITION 659 bp mRNA linear EST 29-MAY-2001
 Chlamydomonas reinhardtii CC-1690, normalized, Lambda Zap II
 ACCESSION BG858032
 VERSION BG858032.1 GI:14239216
 KEYWORDS EST.
 SOURCE Chlamydomonas reinhardtii.
 ORGANISM Chlamydomonas reinhardtii.
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadales; Chlamydomonas.
 1 (bases 1 to 659)
 Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
 McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
 Analyses of the Chlamydomonas reinhardtii Genome: A Model,
 Unicellular System for Analyzing Gene Function and Regulation in
 Vascular Plants; project phase 2
 Unpublished (2000)
 Contact: Charles Hauser
 DCMB Box 91000
 Duke University
 Durham, NC 27708-1000
 Tel: 919 613 8159
 Fax: 919 613 8177
 Email: chauser@duke.edu.

FEATURES
 source
 1..659
 /organism="Chlamydomonas reinhardtii"
 /strain="CC-1690 wild type mt+ 21gr"
 /db_xref="taxon:3055"
 /clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap II"
 /note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO₂ and HS medium bubbled with 5% CO₂. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."
 BASE COUNT 133 a 213 c 201 g 112 t
 ORIGIN

Query Match 7.6%; Score 125.4; DB 12; Length 659;
 Best Local Similarity 53.7%; Pred. No. 2.7e-22;
 Matches 298; Conservative 0; Mismatches 236; Indels 21; Gaps 1;

QY 670 GCNACGGCGGTTGGCTGAGATCAGCGAGATGATCCATGCGAGCATCACTCCTCCACGACGA 729
 Db 125 GCACCGCGGCACTTGGGAGATTAACCGAGATGATTCACCGCGGAGCTTGTGACGACGCA 184
 QY 730 CGTTATCGACGCTTCCGAGACGAGAAACGACCATCCGGAACACGAGCATTCGGAAA 789
 Db 185 CGTGCTCGAGAGTGCACATCCGGCGAGGAGAGACCATCAACAGCATGTATGGCAC 244
 QY 790 CAAAGATGGCGATTTTGGCTGGTGATTTCTTGTGGACGGGCGCTCTGTGTCATTTGGCGAG 849
 Db 245 CGCGTGGCGGCTGCTGGCGGGGACTTCTGTGTCGCCAGTCTCTCGTGGTTCCTCGCCAA 304
 QY 850 GTTGGCCAATCCGAGGTGATTTGAGCTTTTGGCTACTGTTATTGCAAACTTGGTTGAGGG 909
 Db 305 CTGGGCAACCTGGAGGTCAATCAAGCTGATCAGTCAAGTGATTTGCGGACTTTGCCAACGG 364
 QY 910 AGAGTTTCATGCACTTGAATAATACTGTTGATGATGCGATTGAGGCTACGGCGACGACGA 969
 Db 365 CGAGATCAGCCAGCGCGCCAGCCTGTTTGACACCGACAT----- 403
 QY 970 AAGTTTCGATTACTATTTCGAGAAGACTTACTTTGAAGACTGCGTCTTGTGATTCGCAAGTC 1029
 Db 404 CACTCTGGAGCAGTACTCGACAAGTCGTTCTACAAGACCGCCTCGCTCATCGCGCCAG 463
 QY 1030 GTGCAGACGAAGTGGCTTCTGGTGGTGTGCTAGCCCTGAGGTTGCTGATGCTGCTTATGC 1089
 Db 464 CTGCGCTCCGCGCGCTGTTTCTCAGCGACAGCCCTGTGGAGGTGAAGGAGGCCATGTACGC 523
 QY 1090 TTACGGAAGGAACCTTGGTTTGGCATTCAGATCGTCGAGACATGCTCGACTACACCGT 1149
 Db 524 CTACGCAAGCACCTGGGCTGGCATTCAGGTGGTGAGACATCTTGACACTTCACGCA 583
 QY 1150 CTCGCTACCGACCTCGGTAAGCCCGCCGCTGCAGACCTCCAGCTCGGTCTCGCCACCGC 1209
 Db 584 AACACGCGAGCAGCTGGGCAAGCGCGAGGCGCAGGACCTTGCACGCGGTAACCTCACGGC 643
 QY 1210 GCGGCGCTCTTCGC 1224
 Db 644 GCCCGTCATCTTCGC 658

Search completed: January 16, 2003, 08:51:38
 Job time : 1707 secs

THIS PAGE BLANK (USPTO)